BUS. ADM.
LIBRARY

DEC 21 1949

Journal of the
AMERICAN STATISTICAL
ASSOCIATION

DECEMBER 1949

Statistics of the Kinsey Report — *W. Allen Wallis* — 463

The City Block as a Unit for Recording and Analyzing Urban Data — *Edward B. Olds* — 485

The Relation of the Net Reproduction Rate to Other Fertility Measures — *T. J. Woofter* — 501

On Estimating the Mean and Standard Deviation of Truncated Normal Distributions — *A. C. Cohen, Jr.* — 518

On Some Mathematical Problems Arising in the Development of Mendelian Genetics — *Hilda Geiringer* — 526

The Fitting of Logistic Curves by Means of a Nomograph — *Eugene A. Rasor* — 548

On the Best Choice of Sample Sizes for a t-Test when the Ratio of Variances is Known — *John E. Walsh* — 554

Note on Some Errors in "The Evidence of Periodicity in Short Time Series" — *Armen A. Alchian* — 559

William Lane Austin (1871-1949)
James Clyde Capt (1888-1949) — *Stuart A. Rice* — 565

Index to Volume 44, 1949 — 583

Book Reviews by Edmund Churchill, W. Edwards Deming, H. A. Freeman, Sebastian B. Littauer, Walter H. Meyer, Frederick Mosteller, Paul Peach, John W. Tukey — 567

VOLUME 44 NUMBER 248

CHAPTER SECRETARIES

ALBANY	Murray Dorkin, *40 Glenwood Street, Albany, New York*
ASSOCIACIÓN CUBANA DE ESTADISTICA	Hugo Vivo, *Calle 28 No. 270, Apt. B, entre 21y 23, Vedado, Habana, Cuba*
AUSTIN	John R. Stockton, *113 Waggener Hall, University of Texas. Austin. Texas*
BOSTON	Wilfred S. Lake, *Northeastern University, 360 Huntington Ave., Boston 15, Massachusetts*
CENTRAL INDIANA	Miss Lucy Kantz, *1504 N. Penn. #5, Indianapolis 2, Indiana*
CENTRAL NEW JERSEY	William Netschert, Jr., *1 Wayside Ave., Lawrence Township. Rt. 101, Trenton, N. J.*
CHICAGO	Charles G. Wright, *Division of Research and Statistics, Federal Reserve Bank of Chicago, 230 S. LaSalle St., Chicago Illinois.*
CINCINNATI	Lee M. Welsh, *Procter & Gamble, Gwynne Bldg., Cincinnati 1. Ohio*
CLEVELAND	Gale R. Ober, Jr., *76 South Main St., Chagrin Falls, Ohio*
COLUMBUS	Mikhail V. Condoide, *188 West 10th Avenue, Columbus 1, Ohio*
CONNECTICUT	David F. Votaw, *Leet Oliver Hall, Yale University, New Haven, Connecticut*
DENVER	Mary Elizabeth Shannon, *Statistical Dept. P. O. Box 960, Denver, Colorado*
DETROIT	Andrew T. Court, *1517 Iroquois, Detroit 11, Michigan*
HARRISBURG STATISTICAL SOCIETY	Harry Hoyle, *1824 North 4th St., Harrisburg, Pennsylvania*
HAWAII	Margaret S Wright, *P. O. Box 1222, Honolulu, T.H.*
ILLINOIS, UNIVERSITY OF	Donald W. Paden, *Department of Economics, University of Illinois, Urbana, Illinois*
ITHACA	Philip J. McCarthy, *New York State School of Industrial and Labor Relations, Cornell University, Ithaca, New York*
LOS ANGELES	William B Rice, *Box 3519, Terminal Annex, Los Angeles, California*
MADISON, WISCONSIN	Philip J. Fox, *403 Sterling Hall, University of Wisconsin, Madison, Wisconsin*
NEW YORK	Herbert Arkin, *City College of New York, 17 Lexington Ave., New York, New York*
NORTH CAROLINA	George Nicholson, *Department of Mathematical Statistics, University of North Carolina*
OKLAHOMA CITY	Robert S. McCauley, *1710 Exchange, Oklahoma City, Oklahoma*
PACIFIC NORTHWEST	Clyde Courtnage, *Accounting Department, Frederick and Nelson, 5th at Pine, Seattle, Washington*
PHILADELPHIA	Kenneth M. Snader, *Federal Reserve Bank of Philadelphia, Philadelphia 1, Pennsylvania*
SACRAMENTO STATISTICAL ASSOCIATION	Marie Vida Ryan, *3551 12th Ave., Sacramento, California*
SAN FRANCISCO	Harry S. Schwartz, *Federal Reserve Bank of San Francisco, San Francisco 20, California*
ST. LOUIS	Kurt Pohlen, *Catholic Hospital Association, 1402 South Grand Boulevard, St. Louis 4, Missouri*
WASHINGTON STATISTICAL SOCIETY	Margaret Gurney, *Bureau of the Census, Washington 25, D. C.*

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

VOLUME 44 DECEMBER 1949 NUMBER 248

ARTICLES

Statistics of the Kinsey Report . . . . . W. ALLEN WALLIS 463

The City Block as a Unit for Recording and Analyzing Urban Data . . . . . EDWARD B. OLDS 485

The Relation of the Net Reproduction Rate to Other Fertility Measures . . . . . T. J. WOOFTER 501

On Estimating the Mean and Standard Deviation of Truncated Normal Distributions. . . . . A. C. COHEN, JR. 518

On Some Mathematical Problems Arising in the Development of Mendelian Genetics . . . . . HILDA GEIRINGER 526

The Fitting of Logistic Curves by Means of a Nomograph . . . . . EUGENE A. RASOR 548

On the Best Choice of Sample Sizes for a t-Test when the Ratio of Variances is Known . . . . . JOHN E. WALSH 554

Note on Some Errors in "The Evidence of Periodicity in Short Time Series" . . . . . ARMEN A. ALCHIAN 559

William Lane Austin (1871–1949)
James Clyde Capt (1888–1949) . . . . . STUART A. RICE 565

Index of *Journal*, Volume 44, 1949 (Nos. 245, 246, 247, 248) 583

Articles, by author 587

Book Reviews, by author 588

BOOK REVIEWS

DAVID, F. N., *Probability Theory for Statistical Methods* . . . . . JOHN W. TUKEY 567

HERDAN, G., *Quality Control by Statistical Methods* . . . PAUL PEACH 569

JOHNSON, PALMER O., *Statistical Methods in Research* . . . . . FREDERICK MOSTELLER 570

MCNEMAR, QUINN, *Psychological Statistics* . . EDMUND CHURCHILL 572

RISSEK, H., *Quality Control in Production: A Machine-Shop Manual on the Statistical Method of Controlling Product Quality During Manufacture*. . . . . H. A. FREEMAN 574

SCHUMACHER, F. X., and CHAPMAN, R. A., *Sampling Methods in Forestry and Range Management, Second Edition* . . . WALTER H. MEYER 575

WIENER, NORBERT, *Cybernetics* . . . . SEBASTIAN B. LITTAUER 577

YATES, FRANK, *Sampling Methods for Censuses and Surveys* . . . . . W. EDWARDS DEMING 580

Index to Volumes 1–34, 1888–1939, may be obtained from the ASA. The JOURNAL is also indexed in the Industrial Arts Index and the Public Affairs Information Service Bulletin.

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Number 248 DECEMBER 1949 *Volume 44*

STATISTICS OF THE KINSEY REPORT*

W. ALLEN WALLIS
University of Chicago

IN A PREFACE that describes my own preconceptions and attitudes in approaching the Kinsey report, Dr. Alan Gregg of the Rockefeller Foundation, which "contributed a major portion of the cost of the program" (p. vii), observes that "the history of science is part of the history of the freedom to observe, to reflect, to experiment, to record, and to bear witness. It has been a perilous and a passionate history indeed, and not yet ended. . . . The findings of Dr. Alfred C. Kinsey and his associates at Indiana University deserve attention for their extent, their thoroughness, and their dispassionate objectivity. . . . Certainly no aspect of human biology in our current civilization stands in more need of scientific knowledge and courageous humility than that of sex. . . . These studies are sincere, objective, and determined explorations of a field manifestly important to education, medicine, government, and the integrity of human conduct generally. They have demanded from Dr. Kinsey and his colleagues very unusual tenacity of purpose, tolerance, analytical competence, social skills, and real courage. I hope that the reader will match the authors with an equal and appropriate measure of cool attention, courageous judgment, and scientific equanimity" (pp. v–vi).

* Alfred C. Kinsey, Wardell B. Pomeroy, and Clyde E. Martin, *Sexual Behavior in the Human Male*. W. B. Saunders Company, Philadelphia and London, 1948. Pp. xv+804. $6.50. The book is commonly referred to simply as "the Kinsey report," since Kinsey initiated the project in 1938 (p. 10), and did about four-sevenths of all the interviewing that had been done up to the date of publication (p. 11). Similarly, it is common to write "Kinsey" where "Kinsey, Pomeroy, and Martin" or "the authors" would be correct. These abbreviated forms of reference are used throughout this paper.

This paper is a revision of one prepared by invitation of the 1948 Program Committee of the American Statistical Association and presented at Cleveland on 29 December 1948. Some of the material was presented to the Society for Social Research at the University of Chicago on 5 April 1948.

The notes have been collected at the end of the paper, since they should be read after reading the whole of the text. They are reasonably self-contained and can be read without reference to the text, but their numbers have been included in brackets in the text to indicate points that are expanded in the notes.

Sexual behavior is an important and interesting subject that has been considered intensively and extensively from nearly every viewpoint except the simple, factual one of what people actually do. Now, however, we have in the Kinsey report an intensive and extensive collection of facts about overt sexual behavior objectively defined. The report is, in short, a statistical study of sexual behavior. It presents figures on the varieties of sexual activity and the frequency with which each occurs. It reveals the variation of sexual behavior for given individuals and among different individuals, and it relates the variation to age, education, occupation, marital status, religion, and urbanization.

A statistical study cannot, of course, encompass the whole of what we think of as sexual behavior. A definition in terms of overt acts—Kinsey's definition is in terms of activities which ordinarily culminate in orgasm [1]—necessarily excludes aspects that some will describe as "the essential nature of sexual activity," the aspect chosen as essential depending on whether the chooser is an artist, biologist, criminologist, dramatist, psychologist, sociologist, theologist, poet, or philosopher, and perhaps depending also on his own sexual background. As Gregg says, "a great mountain may present aspects that are . . . so different that bitter disagreements can arise between those who have watched the mountain, truly and well, through all the seasons, but each from a different quarter" (p. v). Surely *one* of the significant aspects of human sexual behavior is the statistical one: how many do what, and how frequently? This is what Kinsey has undertaken to find out. That it is not the only thing worth knowing about sexual behavior is no ground for criticism; that it is an important thing to know about sexual behavior can hardly be denied.

There have been, to be sure, other statistical studies of sexual behavior. Kinsey describes nineteen (pp. 23–31), some of them excellent [2]. But no other study has been comparable with Kinsey's in the number of individuals included or in the amount of data about each individual. He has collected data from over twelve thousand persons [3], and has covered over five hundred items [4] for each. In contrast, 4600 is the largest number of persons, and 218 the largest number of items, covered in any of the other nineteen studies listed; and 2484 and 116 are fairer comparative figures, for the 4600 persons were studied superficially and the 218 items were coordinated with Kinsey's.

It is not by any means true, however, that Kinsey's conclusions are all based on statistical data. A great many assertions or implications about religious, ethical, sociological, psychological, and philosophical matters are scattered through the book—so many that I got a cumulat-

ing impression that the author is at heart a social reformer [5]. Most of his conclusions, explicit or implicit, about social and moral issues are based not so much on the statistical data "routinely secured in the interviews" as on "supplementary data" secured by other techniques. "These additional data have come from a considerable list of subjects with whom long-time social contacts have been maintained, in some cases for as long as seven and eight years. . . . While these supplementary records have contributed little to the statistical tabulations of data, they have provided a considerable portion of the detail . . . on the psychologic and social concomitants of sexual behavior, particularly in relation to factors which motivate and control the activities" (p. 74) [6]. It would appear, for example, that most of the material dealing with the attitudes of various social classes toward sexual techniques, their patterns of sexual behavior, and the social implications of class variations in sexual behavior (pp. 363–393) must be based on the supplementary data. In fact, much of the most interesting, and at the same time most controversial, material in the book appears to be based on the supplementary data. At least it clearly is *not* based on the statistical data [7]. One reason the many conclusions and interpretations based on the supplementary data have provoked controversy is that the data themselves are not presented. The passage on pp. 73–75, from which the quotation above was taken, seems to be the only account either of the methods of collection or of the data themselves.

The book contains three rather distinct types of material. There is methodological material, covering procedures for securing subjects, methods of obtaining information from subjects, the reliability of the data, and methods of statistical analysis. There is statistical data on the sexual behavior of certain groups of white American males. And there is "supplementary" interpretative material of a sociological or cultural anthropological kind. These three types of material are distinctly different in pertinence and in scientific quality. Only the second pertains primarily to the subject matter described in the book's title, and only the second is a scientific report, in the sense of an attempt to set forth systematically not only conclusions but the evidence on which the conclusions are based and from which other investigators can plan independent programs for verification or extension. The methodological material has some of the aspects of scientific reporting, but it contains in addition many unsubstantiated assertions [8]. The sociological material consists essentially of insights and opinions about sex in our society. Its character is somewhere between intelligent, technically trained, deeply interested, and thoughtful observation, like Barnard's

work on administrative behavior, and systematic but perspicacious anthropological field work, like Redfield's work on the culture of Central American villages [9].

The three types of material are neither differentiated nor integrated in the volume. The methodological material refers not to the "5300" males to whom the statistical results relate, but to all twelve thousand individuals so far studied. Conclusions based on the sociological interpretations or the supplementary data are frequently stated along with those based on the statistical data, and it is frequently difficult to judge what the basis is for a given conclusion. A clearer distinction would not only have added scientific stature to the report but would have eliminated, or at least focused properly, much of the criticism and some of the enthusiasm that has greeted it.

The book is divided into three parts: (I) History and Method, (II) Factors Affecting Sexual Outlet, and (III) Sources of Sexual Outlet. While this organization results in some unavoidable duplication between the second and third parts, it is on the whole worthwhile to be able to find all the material on, say, the relation of age to sexual outlet brought together in a single chapter of Part II. Similarly, it is an advantage to have all the material on, say, marital intercourse brought together in a single chapter of Part III.

When I first examined the volume, paying attention mostly to its fascinating substantive findings and scarcely at all to its methods, I was very favorably impressed indeed. When I diverted my attention to the general methods I began to note shortcomings; but I felt that these were technicalities—mere blemishes on the surface of the monument, which might modify some of the findings in detail but surely would not affect the broad conclusions. After all, many of Kinsey's figures would still be important and interesting even if we had to allow for an error factor as large as two or even three. But when I spent some time studying the statistical methods in detail, I realized that my confidence in the basic significance of the findings cannot be securely buttressed by factual material included in the volume. In fact, it now seems to me that the inadequacies in the statistics are such that it is impossible to say that the book has much value beyond its role in opening a broad and important field.

Instead of emphasizing directly the reasons for these misgivings about Kinsey's statistical techniques, let me bring them out indirectly by describing some of the improvements that I would like to see in later volumes. Kinsey plans several more reports, and there is every indication in this first one that he welcomes and even seeks criticism

that might help him overcome the numerous imperfections in his work, which he is the first to recognize; so we have here, for once, an instance where there is really a good case for "constructive" criticism. While it is not feasible to make specific detailed recommendations about statistical work unless one is in close touch with all the practical circumstances surrounding it, it is nevertheless possible to suggest tentatively directions in which improvement is possible.

It will be convenient to group my suggestions under three headings, relating to the collection, to the presentation, and to the interpretation to the data.

COLLECTION

I will not discuss the problem of determining the sexual behavior of a given subject—that is, the variables selected, the interviewing technique, and the list of questions [10]—but will confine myself to suggestions concerning the selection of individuals.

The question of sample size ought to be thoroughly reconsidered. Kinsey emphasizes that "the chief concern of the . . . study is an understanding of the sexual behavior of each segment of the population, and that it is only secondarily concerned with generalizations for the population as a whole" (p. 82). He defines his segments according to 12 criteria (sex, race, marital status, age, age of adolescence, education, occupation, occupation of parent, religion, religiousness, urbanization, and geographic residence), each having from 2 to 10 categories. This makes 384,912,000 segments if my arithmetic is correct or "nearly two billion" (p. 81) if Kinsey's is [11]; but as Kinsey points out most of these are fortunately non-existent or rare, and actually only 163 (p. 29)—an impressive enough figure—are covered by the data in the book. Over 40 pages is devoted to the problem of how large a sample is necessary in each segment, and the conclusion is "that a sample of 50 has proven adequate for establishing incidence data, that samples of 100 or 200 are fairly adequate for means and medians, and that samples of 300 or more are quite adequate for determining means and medians . . . but . . . smaller samples may still be taken as indicative of results that may be obtained from larger samples" (p. 683). Unfortunately, however, the discussion of necessary sample size is unadulterated nonsense [12], and represents a prodigious waste of effort.

A proper determination of sample size for Kinsey's material will not be simple, though it should be much easier now that he has collected twelve thousand histories than it would have been earlier. The desirable sample size depends on the requisite accuracy of the results, for one

thing. To specify accuracy, as Kinsey does (pp. 83–85, 736), simply by some arbitrary percentage of the true figure is ambiguous for proportions and may be impracticable for averages, so even the specification needs further consideration in later work. It may be that a double-sampling procedure would be appropriate; perhaps the sequential estimation procedure recently developed by Charles Stein would be of real value [13].

In considering sample size for future work, account must be taken of the fact that many of Kinsey's cases represent the same individuals, and are therefore not independent [10]. Another technical difficulty in determining sample size is that some of Kinsey's sampling is cluster sampling—that is, he selects certain "groups" such as sororities, hitchhikers, or mental institutions, and interviews all members of the group—so the individual histories in his sample may not be independent.

I mention these difficulties not to imply that Kinsey should not treat each history as many observations or that he should not use cluster sampling (though I do feel that both techniques require further analysis [14]) but only to indicate that a proper determination of adequate sample sizes is not an entirely simple problem. A further complication is that multiple measurements are involved in each observation, but this is characteristic of sample-size problems.

Consider next the composition of the sample. Kinsey states that it is "valid to extend generalizations" from his samples to the "163 groups on which data are given" (p. 29). Now a *sine qua non* for generalizing from a sample to a population is randomness. But, as Kinsey points out, in a human survey it is impossible to define a population clearly and then produce a sample that can be guaranteed random (pp. 92–93). On the other hand, any batch of data that we do get is a random sample from some population; that is to say, indefinitely many repetitions of the procedure that produced the sample will produce a population. We have two handles to manipulate in generalizing from human samples: one is to define our population as closely as we can and then attempt to approximate randomness in our sampling procedure; the other is to analyze our actual sampling procedure as well as we can and then attempt to describe the actual population to which it relates [15].

In my judgment, Kinsey in his future work should devote a substantially larger part of his resources—which I realize are limited—to attempting to get a good grasp on one or both of these handles. On the basis of this first report, I believe that far more can be done. With respect to actually defining his populations and then sampling them at

random, the following remark is suggestive: "we have a network of connections that could put us into almost any group with which we wished to work, anywhere in the country" (p. 39). This network together wih the wide publicity that the book has received, may reduce the necessity of spending "days and weeks and even some years . . . in acquiring the first acquaintances in a community" (p. 39).

But, whatever success may be attained in this randomization surely will fall short of perfection. At least it falls short of perfection in such relatively simple matters as marketing surveys, income studies, and even presidential polls. So every step in the sampling process should be paralleled by two steps aimed at studying the sampling process. How many refusals are encountered? How does the refusal rate vary among segments? What are the determinable characteristics of refusers in the various segments? It is said that "the restrained histories have, on the whole, been the more difficult to get" (p. 103). What is the evidence? How much harder? Are there trends in the histories with respect to restraint or other characteristics? It is said that many of the subjects have cooperated in order "to obtain information about some item affecting their personal lives, their marriages, their families, friends, or social relations" (p. 37). Are records kept of these questions, and if so are the questions related to patterns of sexual behavior? It is said that "the greatly disturbed type of person who goes to psychiatric clinics has been relatively rare in our sample "(p. 37). What is the evidence? What similar information is there about the personality types included in the sample, and what more can be obtained? Rare relative to what?

I should judge that nearly half the total cost of analysis for a project such as this would be in checking on the population-sample relationship; but it is worth it, for the validity and usefulness of the research depends fully as much on this as on the soundness of the actual measurements.

A final remark with respect to data collection: In future work it is desirable to plan carefully in advance the methodological checks that are to be included, as for example retakes, comparison of spouses, comparison of interviewers, comparison of remote and immediate recall, comparison of earlier with later results by the same interviewer, analysis of intra-cluster correlation, etc. If in these comparisons the factors not directly involved were balanced out, the results would be substantially more meaningful than are those presented in the book. Furthermore, it is well to avoid comparisons in which, as in the investigation of the clusters (pp. 93–102) or the comparison of age groups (pp. 198, 200) one sample is compared with a second which

includes the first; instead, the first should be compared with the remainder. Application of the statistical principles of experimental design would make it possible to carry out the methodological checks both more effectively and more economically.

PRESENTATION

My strongest recommendation about future reports is that they should tell precisely what was done. My strongest complaint against the present volume is that when I study it in any detail I frequently cannot tell what information Kinsey's conclusions are really based on.

In the first place, I hope that the next volume (or at least a mimeographed supplement to it) will give an account of the actual questions used. There are admittedly difficulties in doing this, for the questions are numerous and "have never been standardized" (p. 51) because "the form of each question has varied for the various social levels and for the various types of persons with whom the study has dealt." For example, "sexual vernaculars must be used in interviewing lower level individuals" and "such vernaculars vary considerably among different groups" (p. 52). Nevertheless, the point which each question covers is said to be "strictly defined" (p. 51) and terms such as "petting" and "prostitute" are precisely explained to the subjects. The interpretation of nearly everything in a study like this depends upon the questions, and their absence makes it impossible for other workers independently to test or to supplement Kinsey's data.

In the second place, future reports should show the actual number of histories involved and how they are distributed with respect to all the controls—age, education, occupation, religion, etc. [15] As a matter of fact, I am not even sure how many white males are covered by the present volume [3]. Furthermore, every effort should be made to present the basic data, at least in a supplementary monograph. Such a presentation should show for each question and each segment the responses given. It is not too late to publish such a supplement to the present volume; this would require a table of 163 columns and perhaps 300 rows, and could probably be presented in less than a hundred pages.

Any information given on the basis of all histories so far collected, as in methodological tests, should also give separately the corresponding figures for the histories covered in the specific report. Obviously statements that are true of twelve thousand cases may not be true at all of a specific set of 5300.

In general, I found the tables apparently meaningful while reading

the book for its substantive content; but when I studied them I found many of them confusing and quite a few downright unintelligible [16]. This is intolerable in a statistical report, but it can be avoided easily in future volumes by competent technical editorial work.

The explanations of statistical techniques struck me as thoroughly unclear. The explanation of the Accumulative Incidence Curves, described as "the one new statistical tool which we have had to develop for this study" (p. 114) and its superiority over the ogive, would have left me bewildered had I not come across essentially the same idea clearly explained in the June 1946 *Journal of the American Statistical Association* in an article on "The Operating Life of B-29 Engines" by Oscar Altman and Charles Goor, who casually refer to the technique as a standard actuarial device. And Kinsey's formula for the median (p. 113) is just wrong [17]. Again, these things are intolerable in a statistical report; but they can surely be remedied in future work.

As for the so-called "U. S. Corrections" (pp. 105–109), I have been unable to understand them. The explanation sounds straightforward, but a few calculations I have made with them do not check with Kinsey's [18]. The U. S. Corrections are obviously intended as a set of weights for combining various of the 163 segments into averages applicable to broader groups, and of course some such set of weights is needed.

INTERPRETATION

Most of what I have to suggest under this heading has been covered by implication under the "collection" heading. In general, if a statistical investigation of this kind is well planned and the data properly collected the interpretation will pretty well take care of itself. So-called "high-powered," "refined," or "elaborate" statistical techniques are generally called for when the data are crude and inadequate—exactly the opposite, if I may be permitted an *obiter dictum*, of what crude and inadequate statisticians usually think.

Kinsey's data have very properly been subjected to a minimum of processing. His measurements of incidence, frequencies, and accumulative incidence are straightforward and sound, and no doubt his U. S. Corrections are too, if I only understood them. While he smooths most of his curves, he always shows the original data too [10]. All of this I hope to see continued in later volumes.

The only measure which is seriously mishandled is the range. This is a statistic which can be interpreted only if one takes into account the number of observations, makes stringent assumptions about the

normality or other mathematical characteristics of the population and about the independence of the observations, and has at hand an appropriate set of tables relating sample range to population variability. If the variability in several populations is constant, the range in samples will be larger the larger the sample. Kinsey falls into a fallacy when he shows (p. 234) the next-to-highest frequency of outlet found at various ages and interprets the marked decline with age to indicate that variability declines with age. It would not be surprising if it does in fact decline with age, but the table does not show it without further interpretation, for the number of cases declines from 3,012 in the youngest group to 58 in the oldest. Sample ranges should not be shown in future studies but some other measure of variability should be substituted, for example the standard deviation, the average deviation, or—probably more practicable—the distance between the first and ninth deciles.

It would be desirable in further work to distinguish between differences that are not statistically significant—in other words, that aren't there as far as the data show—and those that are statistically significant but of little consequence. In several instances the present book disregards differences which the data establish, but which if admitted to exist could be shown to be small enough to be negligible for many purposes. For example, different interviewers seem to get significantly different results (as is characteristic of many measuring devices), but the differences are apparently not great enough to affect the conclusions seriously (pp. 133–143).

Terman's criticisms of Kinsey's inerpretation of the data, and also his criticisms of other aspects of the work, all seem to me entirely sound [19]. Terman maintains convincingly that Kinsey has misinterpreted his data when, for example, he concludes that "the sexual patterns of the younger generation are . . . nearly identical with the sexual patterns of the older generation in regard to . . . many types of sexual activity" (p. 397), and when he asserts that sexual patterns are stable throughout life, assuming in childhood the pattern of the occupational group to which the individual will ultimately move (p. 419). But it is hard to make specific recommendations for avoiding this kind of misinterpretation, except to reiterate that a substantially better quality of statistical work is essential if future research is to have value.

After studying the volume I still feel that it is a pioneering and monumental work in an important field. And of the authors' virtues as listed by Dr. Gregg, I still admire their tenacity, their tolerance, their social skills, and their courage; but I have major reservations

about their analytical competence insofar as that means statistical competence. The work will take a place in the history of this subject, it seems to me, analogous to that occupied in the history of price studies by Thorold Rogers' seven volumes on the *History of Agriculture and Prices in England, 1259–1793*, of which the *Encyclopedia of the Social Sciences* says, "Even his severest critics express admiration of his scholarly labors in extracting prices from such sources as the bailiff's accounts of the property held by the colleges of Oxford and Cambridge and the great monastic corporations of the Middle Ages. It is the interpretation of these figures . . . which is open to question and correction." "From the point of view of modern statistical technique these studies leave much to be desired. They have been modified and supplemented by recent works, especially those of Beveridge, Usher, and Hamilton, so that it is now possible to trace with some assurance the general trend of prices in western Europe through the interesting period of the 'price revolution'" [20]. As a consequence of Kinsey's labors it will no doubt be possible at some future date to describe with some assurance the statistical pattern of sexual behavior.

NOTES

[1] The following passage comes as close to defining sexual activity as any I have found in the book: "The sexual activity of an individual may involve a variety of experiences, a portion of which may culminate in the event which is known as orgasm or sexual climax. There are six chief sources of sexual climax. There is self stimulation (masturbation), nocturnal dreaming to the point of climax, heterosexual petting to climax (without intercourse), true heterosexual intercourse, homosexual intercourse, and contact with animals of other species. There are still other possible sources of orgasm, but they are rare and never constitute a significant fraction of the outlet for any large segment of the population" (p. 157). "Outlet," as used in the report, seems to be equivalent to "orgasms" (pp. 193, 683). The six forms of behavior listed constitute the sexual behavior studied in the report; two measurements are presented for each type: (1) incidence, the proportion of individuals engaging in the activity, and (2) frequency, the number of orgasms per week from the activity.

[2] Kinsey tends to criticize other studies on grounds that sometimes amount to nothing more than that their methods differ from his. For example, he asserts that Terman's data on 1,242 married couples "would have been more reliable if they had been obtained by direct interviewing" (p. 31), though he neither gives nor cites evidence bearing on the issue, either in discussing Terman's and other questionnaire studies (p. 31) or in discussing his own technique of direct interviewing (Chap. 2). On p. 11 he does assert that "during the first year the value of personal interviewing as opposed to the questionnaire technique was subjected to some testing," but I have found no other reference to this testing or its results. Again, on p. 42 it is asserted, in

effect, that "things . . . can be done in a person-to-person, guided interview . . . that can never be done through a written questionnaire, or even through a directed interview in which the questions are formalized and the confines of the investigation strictly limited." Despite the lack of supporting evidence, this seems plausible to me; but so would it if its meaning were reversed by interchanging "person-to-person guided interviews" with "written questionnaire" or with "directed interview," and so does the following, from p. 62: "Whether the techniques which have been used in the present study would be equally effective with other persons engaged in studying other problems, is a question which must be answered empirically by each investigator in connection with his own special problems."

[3] The number of individuals involved in the study is given on p. 10 as 12,214. On p. 5, however, there is an outline map of the United States with the legend "Sources of histories. One dot represents 50 cases." The map contains 427 dots, so presumably represents 21,350 cases. Even if for each state one of the dots represents fewer than fifty cases, 427 dots would represent at least 19,000 cases. Kinsey uses "histories" and "cases" with different meanings, so their apparent equivalence on this map is confusing: "history" refers to the data for one individual, and "case" to the data for one individual during one five-year period of his life. The largest number that I have seen mentioned in the book as the total number of cases is 14,084 (p. 220). These are presumably derived from the 5,300 white males to whom the data in the book are said to relate (pp. vii, 6). The ratio of cases to histories for the white males thus appears to be 14,084:5,300 or 2.66. Application of this ratio to all 12,214 histories suggests over 32,000 cases, which is as much too high as the number of histories is too low to account for the 427 dots on the map.

Actually, it is not quite clear that the total number of histories is 12,214, as stated on p. 10. The histories are said to include 5,300 white males, about 1,000 non-white males (p. 6), and 5,800 females (p. 29). The discrepancy of 114 is presumably due to rounding to hundreds, although the total number of males is given at least once as "about 6,300" (p. 6) and at least once as "6,200" (p. 29). However, 12,214 is the total number shown both in the table distributing histories by year of collection (p. 10) and in the one distributing them by interviewers (p. 11).

That 5,300 is the number of "white males who have provided the data for the present publication" (p. 6) is not confirmed by any of the statistical tables in the book. The largest total I have noticed (often the totals are not shown but have to be computed) in any of the tables that appear to cover all of the white males is the 4,120 shown distributed by religion in Table 41, p. 208. This same table shows 4,940 males distributed by occupation, but since the adjacent column which is said to distribute 179 males by occupation totals 237, it may be that some individuals are classified under more than one occupation. Indeed, it may be that the 4,120 include some classified under more than one religion, for 4,102 is the number in the same table classified by education, and 4,069 is the number classified by age at onset of adolescence. Tables 37 and 41 both include what appears to be the same distribution by age at onset of adolescence, but the frequencies differ: Table 37 includes 521 more males (about one-eighth more) than Table 41, and these additional 521 have a distinctly different distribution by age of adoles-

cence than the 4,069, relatively more (37.0 per cent instead of 30.8 per cent) being below 13 years and relatively fewer (30.7 per cent instead of 36.3 per cent) above 13. Table 67, p. 298, shows 4,606 males distributed by age at onset of adolescence through 16 years, which is 43 or 44 more than shown in Table 37 for the same ages; adding to these 4,606 the 28 shown in Table 37 as reaching adolescence after 16 would give a total of 4,634 males. Table 36, p. 186 shows only 3,730 cases distributed by school grade at adolescence, and Table 35, p. 184, shows totals ranging from 1,355 to 3,573 cases for five distributions by age of occurrence of various developments in adolescence. The number for whom information is not available on a given item is never shown. In general, very little is revealed in the statistical data about the number of males covered in the volume.

The following "definition" of cases is said to "have been applied . . . throughout the present volume" (p. 682): "*Cases.* Showing the size of the population on which the data in the tables are based" (p. 683). Table 44, p. 220, shows 14,084 cases distributed by age from adolescence to 85, of whom 11,467 are 30 or under. Table 40, p. 198, shows 14,083 as the number of cases for "all ages, adolescence to 85," also with 11,467 of them 30 or under. Table 41, p. 208, however, shows the number of cases from adolescence through 30 as 11,985; so perhaps 518 cases should be added to the 14,084 in trying to discover the total number of cases involved in the study. Tables 104 and 105 (pp. 410 and 412) both show 13,359 as the total number of cases, 11,314 single and 2,045 married; 9,286 aged under 33 and 4,073 aged 33 or over. (Incidentally, this is one of the few pairs of tables I have had occasion to compare that have not proved inconsistent in their totals; the pair on pages 10 and 11 is another exception, and Tables 40 and 44 miss by only one case. Tables 104 and 105, though consistent with one another, show fewer cases 32 years of age and under than are shown as 30 and under in Tables 40, 41, and 44.) Tables 152–154, pp. 686–734, the so-called "Clinical Tables," include 15,746 cases (11,725 single, 3,275 married and 746 previously married), according to my addition. The numbers of cases shown in these clinical tables are hard to reconcile with one another, however, for the sum of the numbers shown in various subdivisions sometimes exceeds, and sometimes is exceeded by, the number shown for the whole group. For example, the data on pp. 688–690 for single white males, age group adolescence through 15 years, educational level 13+, show the whole group as 2,799 cases; but the sum of corresponding figures for the urban (2,587) and the rural (352) subdivisions of the group is 2,939, and the sum of the corresponding figures for the six religious subdivisions of the same group is 2,974. Such a result, where the whole is less than the sum of its parts, could occur if, for example, an individual who shifted from one subdivision to another were classified in both for the five-year interval in which the shift occurred—though this practice would be unsound. On the other hand, the whole frequently turns out to exceed the sum of its parts; for example, for the same age, color, and marital status just discussed but educational level 9–12, the total number of cases for the whole group is 606, but the urban (459) and rural (124) figures total only 583.

The term "population" in the definition of cases quoted above seems actually to mean "sample." The rest of the paragraph of which the quotation is

the first sentence deals with the adequacy of samples of various sizes. In general, Kinsey seems to be aware of the distinction between the statistical concepts of "sample" and "population" and of the fundamental importance of the distinction (see, for example, the section on "The Taxonomic Approach," pp. 16–21), but he misuses the terms in a way that is frequently disconcerting. For example, the heading "population in sample" appears in various tables (e.g., p. 190) and "sample population" in others (p. 188). Sometimes the phrases "sample population" and "U. S. population" (p. 188) seem to represent the sample-population distinction. The phrase "the whole population involved in the present study" (p. 194) seems to mean "the whole sample."

"Educational level 13+" is referred to above in discussing the numbers of cases shown in Tables 152–154. To Kinsey "13+" means "ultimately more than 12 years," i.e., that at least a start has been *or ultimately will be* made to college. In using the tables as a basis for comparison of any particular male, it is necessary either to exclude "persons who are still in school, since there is no certainty how far they will go before they finally terminate their education" (p. 331), or to "predict, on the basis of his home background, the amount of his future schooling" (p. 682); and presumably similar requirements apply to those who have left school before reaching the highest educational level and who might later resume their education.

It is doubtful that the age division referred to above in connection with Tables 104 and 105, and also involved in Tables 98–103, is between "persons who were 33 years of age or older at the time they contributed their histories" and "persons who were younger than 33 at the time of contributing," as stated on p. 395, since Tables 98–100 include data on the behavior of the younger group at age 33. Perhaps the younger group was actually 33 and under and the older group over 33.

[4] Five hundred and twenty-one is said to be the number of "items which are systematically covered on each of the histories in the present study" (p. 32), but this is perhaps an exaggeration, partly because the count depends on what is considered an item and what is considered merely one of several possible responses for an item, and partly because for any individual many of the items are inapplicable.

The list of "Items Covered on Sex Histories" (pp. 63–70) shows nine major groupings under which are a total of 71 numbered headings. The items listed under some headings call for distinct pieces of information; for example, under "educational history" are listed years of schooling, colleges attended, college majors, age upon leaving school, and age while in high school. The items listed under other headings, on the other hand, constitute little more than a list of possible answers to a single question; for example, under "recreational interests" are listed, among others, moving pictures, dancing, cards, hunting, fishing, reading, sewing, music, sports. There is no way to tell whether the claim of 521 includes items like those under recreation on a par with those under education.

The "actual number of items covered in each case" is described on p. 63 as "usually nearer 300, and the number involved in the histories of younger and less experienced individuals is often less than that." This suggests that 300 *or less* is the number of items involved for any one person. The following

passage from p. 50 suggests, however, that 300 *or more* is the number for each individual: "On each history in the present study there has been a systematic coverage of a basic minimum of about 300 items. This minimum is expanded for persons who have extended experience. . . . The maximum history covers 521 items." On the other hand, it appears from p. 51 that numerous items beyond the list of 521 may be covered in some histories, e.g., males with elaborate techniques of masturbation, individuals who have had some complex relation with their parents, identical twins, highly intelligent individuals with considerable experience in a socially taboo type of behavior, individuals involved in masochism or sadism, persons who are handicapped, have lived in foreign countries, have had experience in military groups, etc. "As scientific explorers, we, in the present study, have been unlimited in our search to find out what people do sexually" (p. 51).

Kinsey's comparison of his study with nineteen others in regard to number of items covered is a little misleading (pp. 28–29). For example, Ramsey's 218 items are described as 41.9 per cent as many as Kinsey's; but for boys from Junior High Schools, Y.M.C.A.'s, and Boys' Clubs, the group covered by Ramsey, Kinsey's list of relevant questions may be not much larger than 218—certainly it is not nearly two and one half times as large, as is implied by the 41.9 per cent figure. Thus what at first appears to be a difference in the amount of information for each individual is at least partly only a reflection of the difference in the number and variety of individuals covered. In fact, we learn on p. 30 that Ramsey's work was "based on personal interviews which were coordinated with the list of questions and the techniques of the present study."

[5] If we accept the implications of Kinsey's assertions (p. 199) that "there is, inevitably, some correlation between these rates [orgasms per week] and the positions which these persons take in a public debate" on sex instruction and administrative policies of educational institutions, that "the policies that ultimately come out of such meetings [on juvenile delinquency, law enforcement, and sex laws] would reflect the attitudes and sexual experience of the most vocal members of the group, rather than an intelligently thought-out program established on objectively accumulated data," and that "often the conclusions [of scientific discussions of sex] are limited by the personal experience of the author," then many of his own conclusions and implications—especially those not based on the statistical data—require for their evaluation some indication of where he himself falls in his various distributions. Acceptance of this line of argument might "explain away" many of Kinsey's social and ethical judgments, as, for example, if he were one of the six persons described on p. 217 as having the highest long-time averages.

[6] The first omission from this quotation describes in more detail the social contacts used to collect the supplementary data: visits to the subjects' homes; visits with them to their friends' homes, theatres, concerts, night clubs, taverns, and other places of recreation; correspondence; records of their sexual activities; photographs of their drawings; transcripts of their court, institutional, or social agency records. This impressive account of Kinsey's own collection of supplementary data should be contrasted with the following: "Sometimes social scientists hobnob as tourists in some social milieu sufficiently removed from their own to make it possible for them to

acquire 'impressions' and 'hunches' about 'social patterns' and 'motivations of behavior' in whole cultures. . . . The day seems overdue when scientists studying human material will forsake barbershop techniques . . . " (p. 19). This brings to mind Alfred Marshall's "general rule that in discussions on method and scope, a man is nearly sure to be right when affirming the usefulness of his own procedure, and wrong when denying that of others" (*Principles of Economics*, Eighth Edition, p. 771). Another example is provided by comparing the passage on p. 201 objecting to conclusions about sexual behavior based on persons who come to a clinic, with the passage on p. 37 asserting that while many of Kinsey's subjects were obtained because of public knowledge of the project as a source of help with personal sexual problems, these were the everyday sexual problems of the average individual.

[7] Actually, it is hard to see how even the supplementary data could support such apparently quantitative assertions as "most of the tragedies that develop out of sexual activities are products of this conflict between the attitudes of different social levels" (p. 385), or "most of the complications which are observable in sexual histories are the result of society's reactions when it obtains knowledge of an individual's behavior, or the individual's fear of how society would react if he were discovered" (p. 202).

[8] Many of Kinsey's methodological assertions are of considerable interest, so it is regrettable that he has not indicated their basis. Examples are: "Appreciation must be sincere, else it will not work" (p. 37). "The underworld requires only a gesture of honest friendship before it is ready to admit one as a friend, and to give histories 'because you are my friend' " (p. 36). "The experienced interviewer knows when he has established a sufficient rapport to obtain an honest record, in the same way that the subject knows that he can give that honest record to the interviewer" (p. 43). "People understand each other when they look directly at each other" (p. 48). "We attempted to follow standard practice [of making records only after the subject has left at the close of an interview] early in this study and found that it introduced a tremendous error into the records" (p. 50). "When one is dealing with such a socially involved question as sex it becomes particularly important to ask direct questions. . . . Euphemisms should not be used . . . " (p. 53). " . . . we always begin by asking *when* they first engaged in such activity. . . . It might be thought that this approach would bias the answer, but there is no indication that we get false admissions . . . " (p. 53). "Looking an individual squarely in the eye, and firing questions at him with maximum speed, are two of the best guarantees against exaggeration" (p. 54). Also, see Notes 2 and 15.

The correctness of these assertions is not in question here. They are cited simply as illustrations of propositions asserted without the evidence necessary to enable another investigator to evaluate them. In this sense they differ from Kinsey's conclusions about male sexual behavior, insofar as these are accompanied by statistical data whose source is explained.

[9] See Chester I. Barnard, *The Functions of the Executive* (Harvard University Press, 1938, pp. xvi+334) and *Organization and Management* (Harvard University Press, 1948, pp. xi+244); Robert Redfield, *Tepoztlan, A Mexican Village: A Study of Folk Life* (University of Chicago Press, 1930, pp. xi

+247) and *The Folk Culture of Yucatan* (University of Chicago Press, 1941, pp. xxiii+416).

In quality, as contrasted with type, Kinsey's work is not comparable with Barnard's or Redfield's. Barnard, for example, seems to have a deeper understanding of the inconsistency between legalistic statements of the principles of administrative authority in a social organization and actual behavior in the same organization, than Kinsey does of the inconsistency between our "publicly pretended code of morals" (p. 197) or our "socially pretended custom" (p. 203) and his finding that the persons involved in "illicit activities, each performance of which is punishable as a crime under the law, . . . constitute more than 95 per cent of the total male population" (p. 392).

Incidentally, this 95 per cent figure needs to be interpreted with more care than Kinsey uses. It means that 95 per cent of all white males either have engaged at least once in their lives in an "illicit" activity or can be expected to engage in one at least once, if they live to be 85. (Actually the same figure results if we assume only that they will live to be 45.) Several of the statements Kinsey makes in connection with this 95 per cent figure imply that 95 per cent of the total male population *has* engaged in illicit activities. This is logically equivalent to interpreting "all males now living will ultimately die" to mean "all males now living have already died."

No source is cited for the figure 95 per cent and it cannot be verified from the statistical data given in the book. Even if it is correct and were correctly interpreted, to conclude from it that "only a relatively small proportion of the males who are sent to penal institutions for sex offenses have been involved in behavior which is materially different from the behavior of most of the males in the population" (p. 392) would be a non-sequitur. As Terman says, "it is as though one said that if 95 per cent of all males have at some time in their lives stolen something, those who are sent to penal institutions for theft or burglary are not materially different from most males in the population." (Lewis M. Terman, "Kinsey's 'Sexual Behavior in the Human Male': Some Comments and Criticisms," *Psychological Bulletin*, vol. 45, 1948, pp. 443–459; the quotation is from p. 456.)

Kinsey makes a similar statement about adolescents: "On a specific calculation of our data, it may be stated that at least 85 per cent of the younger male population could be convicted as sex offenders if law enforcement officials were as efficient as most people expect them to be. The stray boy who is caught and brought before a court may not be different from most of his fellows, but the public, not knowing of the near universality of adolescent sexual activity, heaps the penalty for the whole group upon the shoulders of the one boy who happens to be apprehended" (p. 224).

[10] Terman, in the review cited in Note 9, makes a number of sound criticisms of the methods of obtaining data from a given subject. One that must receive attention in any consideration of Kinsey's statistics is his use of data based on long-distance memory. Each individual is asked about his activities as far back as he can recall, and his reports are tabulated by five year age intervals. Thus each individual provides several cases. "In the computation of mean frequency of masturbation at age 15, for example, the memory report of a 50-year-old counts as heavily as the report of a 15-year-old" (Terman,

p. 446). One important consequence of this, aside from inaccuracies that it may introduce, is that many of the "cases" on which the statistical analysis is based are not independent. This might explain, for example, at least in part, Kinsey's finding that patterns of sexual behavior are remarkably constant throughout life, assuming in early childhood the pattern of the occupational and educational level to which the individual ultimately will belong (p. 419). It also invalidates, at least partially, Kinsey's notion that "smooth trends in such curves are evidence of their approach to reality" (p. 132), for most of the smooth trends shown are age trends and therefore involve many of the same individuals in the successive age groups. In fact most of the charts to which Kinsey has added smoothed curves are accumulative incidence curves; not only are the successive age groups not independent here, but the calculations for any age group involve cumulating cases over other age groups.

Incidentally, the fact that some 62 of the 173 charts contain smoothed curves is surprising in view of the statement that "All of the frequency curves in this volume are based on the actual calculations, and in no instance have they been smoothed by any process or approximated by interpolations or other sorts of estimates or predictions" (p. 111). Most of the 62 smoothed curves are accumulative incidence curves, which show per cent of total population having had a specified type of experience—a sort of cumulative frequency adjusted for exposure-to-risk—plotted against age, so perhaps are not regarded as "frequency" curves. Figures 36 and 37, however, both show frequency data that have been smoothed by cumulating and further by use of a curve. Personally, however, I have no complaint against the smoothed curves, since all charts show the actual observations clearly. The simple frequency polygons, showing per cent of cases in various class intervals by average number of orgasms per week, are nearly all too erratic to permit of useful smooth curves; they thus fail to provide whatever confidence in the data would be lent by smoothness.

[11] The criteria by which cases are classified and their numbers of categories are: sex, 2; race-cultural group, 11; marital status, 3; age, 18; age at adolescence, 6; educational level, 9; occupational class, 10; occupational class of parent, 10; rural-urban background, 5; religious group, 3; religious adherence, 4; geographic origin, unspecified. The product of the 11 category numbers is 384,912,000. If Kinsey intends to represent geographic origin by 48 states, which are the only geographic units he mentions, there will be nearly 20 billion categories. So far, he has made no classifications by geographic origin, but he says "state of residence for the most continuous period of time, and the place of residence during the childhood and adolescent years, will probably represent the most significant part of the data" (p. 81). It seems to me that a few broad regions might suffice; five regions would result in "nearly two billion" segments.

[12] Quinn McNemar cites the following four fallacies in Kinsey's efforts to determine a proper size of sample: "(1) Failure to recognize the fact that the sampling stabilities of means, medians, and modes are not a function of their magnitudes, but rather of trait variability. . . . (2) Failure to consider the fact that these statistics differ markedly from each other in their sampling errors. . . . (3) Failure to observe the fact that the sampling stability of

percentages is not a linear function of their magnitudes but rather of their degree of remoteness from 50 per cent. . . . (4) Failure to note that convergence of sub-sample values to total group values must be more rapid when sub-samples are drawn from small (finite) groups than when drawn from larger groups. . . . In brief, incognizance of four elementary statistical principles renders worthless this elaborate effort to determine how large N should be for a sub-group" (quoted on pp. 450–451 of the Terman review cited in Note 9; the positive part of McNemar's third point seems to be an imprecise allusion to the fact that the sampling variance of a proportion is a linear function of the squared difference between the proportion and one-half.)

Such incompetence on this and other statistical points is surprising in view of Kinsey's statement, widely cited among statisticians, that "the statistical set-up of the research was originally checked by Dr. Lowell Reed of the School of Hygiene and Public Health at The Johns Hopkins University. A long list of persons experienced in sampling and in other aspects of statistics has been constantly available for consultation" (p. viii). At the Cleveland meeting where this paper was presented, Helen M. Walker, who presided, read a letter from Reed to Kinsey dated December 10, 1948 which said, in part, "I have been troubled at the flood of criticism that has been leveled at your work by the statisticians, mainly because I know that there is value in your work, but secondarily because you included my name in the preface with the implication that I had some responsibility for the analysis. As you of course remember, I saw your work only on the occasion of a two-day visit to Bloomington in December, 1942. On the basis of that visit, I joined heartily with the Committee in recommending to the National Research Council that support be continued, but a part of the recommendation was that appropriate arrangements within the budget should be made to strengthen the project from the statistical side. I became so busy with work connected with the war that I lost all contact with the project and I don't know what was done, if anything, to carry out this recommendation. If the type of statistical guidance had been provided that I had in mind, I feel sure that you would now be free of some of the criticism that is now being justly leveled at the work. . . . As you know, . . . I have seen nothing of the work between that two-day visit in 1942 and the public appearance of the book." (Transcribed from recording made at the meeting by Chester I. Bliss.) As for the "list of persons experienced in sampling and in other aspects of statistics," it would appear that either they were not consulted or else their advice was not followed.

[13] Stein's sequential estimation procedure was presented before the Institute of Mathematical Statistics at Madison on 10 September 1948, but has not otherwise been published. It is a method of determining, to prescribed accuracy with prescribed confidence, the mean of a normal distribution whose standard deviation is unknown. On the average, it requires hardly any more observations than would be required in standard single sampling if the standard deviation were known.

[14] It would be interesting to have the data for the cases in a given age group classified by the age of the subject at the time of the interview. Any systematic variation with age at interview could, however, be interpreted as reflecting either memory effects or time trends; but independence between

behavior reported for a given age and age at interview would be necessary to confirm both Kinsey's handling of cases and his conclusion that generation-to-generation changes are minor (pp. 394–417). Similarly, a study of the intra-class correlation in clusters of the kind Kinsey uses would be valuable.

[15] Kinsey's discussion of his sample is inadequate in two crucial respects: he does not explain adequately his procedure of selecting cases, and he does not describe adequately the determinable characteristics of the sample that he got.

With respect to sampling procedures, he says little except that randomness is difficult to achieve. He does not suggest that he made any efforts to approximate it; instead, he says "since it is impossible to secure a strictly randomized sample, the best substitute is to secure one hundred per cent of the persons in each social unit from which the sample is drawn" (p. 93). No basis for the assertion is given, but probably the hope is to include not only the most willing but also the more reluctant subjects. Whatever gain there might be in this respect, however, would be offset at least partially by the fact that efforts to get all of a group were not ordinarily made unless at least half had already been obtained (p. 95); thus, those groups containing relatively many reluctant subjects would be less likely to be the objects of hundred per cent drives than those containing relatively few.

As to the composition of the sample actually secured, it is hard to learn much about this; even its size is uncertain (see Note 3). Scattered through the book are various scraps of information about special groups that have been sampled. Terman, on p. 447 of the review cited in Note 9, lists some of these: Of the 62 hundred-per cent groups, 42 were of college level and 7 were delinquents or inmates of penal and mental institutions (p. 95.) "Perhaps half" of the histories were obtained through contacts resulting from lectures (p. 38). Seventeen penal or correction institutions have provided histories (p. 15). Five underworld communities and five homosexual communities are represented—they are listed as "social or civic organizations" (p. 16). There are data on 1,200 persons convicted of sex offenses (p. 392). In addition, a passage on p. 38 strongly suggests to me that "several hundred psychoanalysts, psychiatrists, physicians, clinical psychologists, social workers, and other professional persons [who] have had an especial interest in observing the interviewing techniques" were included; and this interpretation is reinforced by the section on "The Confidence of the Record" (pp. 44–47). On pp. 14–15 we learn that the sample includes persons who have been students at 528 colleges, and that 14 of these have contributed 100 or more histories apiece; even assuming that each of the 14 has contributed only 100 and that each of the other 514 has contributed only one, this accounts for 1,914 histories, or nearly 16 per cent of the total number (12,214). The map on p. 5 showing the source of the data (see Note 3) has 191, or 45 per cent, of its dots in Indiana and the four adjoining states, Illinois, Kentucky, Michigan, and Ohio—five states which contain about 20 per cent of the U. S. population (19.4 per cent in 1946).

Table 41, p. 208, which was discussed in Note 3 above, comes as close as any to revealing the characteristics of the sample. It suggests that about 60 per cent of the histories are college level (in fact, that 26 per cent have

training beyond college); that 1.6 per cent are from the underworld, 0.5 per cent are business executives, and 61 per cent are white collar or professional workers; that 76 per cent are Protestants, 12 per cent Catholics, and 12 per cent Jews; that 74 per cent are religiously inactive. What we clearly should have, however, is a definite statement of the distribution of the histories among the 163 segments for which conclusions are drawn.

[16] An example of a table that is difficult to understand is Table 14, "Comparisons of data obtained from spouses" (p. 126). The second column in this table is headed "items involved" and the third column is headed "unit of measurement." The seventh and eighth columns are headed "mean of husband's reports" and "mean of wife's reports." For the item "pre-marital acquaint." the unit of measurement is "12 mon."and the means of husband's and wife's reports are 42.11 and 40.88. For "engagement" the unit is "4 mon." and the means are 12.64 and 12.85. For "lapse, marr.—first birth" the unit is "6 mon." and the means are 28.05 and 28.19. These figures seem to say that the couples in the comparison were, on the average, acquainted 41 to 42 years before marriage, engaged over 4 years, and married 14 years before their first child was born! The data seem more plausible, however, if we assume that the units of measurement are one month in all three cases. The "units" given seem only to describe the amount of discrepancy which, if not exceeded, is called zero, i.e., identical response for both spouses. The fourth column is headed "ident. rspns. %."

Tables 152–154 are puzzling because, as mentioned in Note 3, the numbers of cases in subgroups add up sometimes to more and sometimes to less than the number of cases shown for the whole group.

[17] Kinsey says that the formula for the median is $(n+1)/2$ (p. 113). The text following this formula, however, gives a correct explanation of the median. With respect to the arithmetic mean, his formula (p. 112) is correct, except that he does not define its symbols, but it is followed by three erroneous assertions: first, that "a mean represents the total number of measurements . . . in each group divided by the number of individuals in the group"—he means, of course not the total number but the sum of the measurements, and an illustrative example included parenthetically at the omitted part of the quotation is correctly handled; second, that "the mean represents the midpoint of the measurements"—the median is the midpoint either in the sense that as many observations lie above as below it, and the mean in the sense that the sum of the deviations is as great above it as below it, but neither the mean nor the median is midway between the extremes; third, that the mean's "position . . . is therefore [i.e., because it "represents the midpoints of the measurements"] materially affected by the presence of even a few high-rating individuals in a population [i.e., sample—see Note 3]; and . . . a few high-rating individuals affect the means more than a large population [i.e. number] of low-rating individuals"—which suggest that Kinsey really believes the second of his erroneous assertions.

Kinsey also asserts that "where most of the individuals in a sample belong in a frequency class which is midway between the extremes of the distribution, and where an equal number of individuals lie in symmetrical distribution on either side of the midpoint, the mean becomes identical with the median" (p. 113). It is not clear whether this means that *either* of the two

conditions will bring about the identity, or that *both* are necessary to bring it about. Actually, the first condition is irrelevant; and the second condition is sufficient but not necessary for the identity.

[18] As a simple illustration of the use of the "U. S. Corrections," consider Table 29, "Continuity of pre-adolescent sex play with adolescent activity" (p. 174). The first section of this table shows "% with continuity" for three educational levels, as follows: level 0–8, 77.0; level 9–12, 67.4; level 13+, 29.8. Since the sample sizes ("cases") for these levels are 243, 221, and 763, the per cent for all three levels pooled into a single sample of 1,227 would be 45.9. Obviously, however, an over-all average should weight the three levels not in proportion to their sample sizes as pooling does, but in proportion to their population sizes. These population weights are given for 1940 as 53.21, 35.13, and 10.39, respectively, with 1.2 per cent of the population not reporting educational level (Table 11, p. 108). Combining the three observed percentages with these weights gives an over-all percentage of 68.6. Kinsey shows 64.9, however. Similarly, for the second and third sections of the table my calculations give 60.7 and 44.8, but Kinsey shows 54.7 and 42.1. In a similar check of the three largest percentages shown in the final column of Table 38 (p. 190), I find 68.24, 14.42, and 11.29 where Kinsey shows 68.39, 12.53, and 13.11. Since Kinsey assures us that "all mathematical calculations on this project have been performed twice, independently by each of two persons" (p. 109), I assume that I have not understood the "U. S. Corrections" correctly.

A curious feature of the "Tables for U. S. Corrections" (pp. 106–108) is that the age groups for which the weights are shown are not the same as those used throughout the report, but are a year lower, e.g., 15–19 and 20–24 instead of 16–20 and 21–25. This does not affect the discrepancies mentioned in the preceding paragraph, however.

[19] Terman's review is cited more specifically in Note 9. Another review which makes several sound criticisms of the interpretation is that by Jacob Goldstein and Nicholas Pastore, "Sexual Behavior of the American Male: A Special Review of the Kinsey Report," *Journal of Psychology*, 26 (1948), pp. 347–362. Both of these reviews have been helpful in preparing this paper, which has also benefited from critical readings by Milton Friedman and by L. J. Savage.

[20] The quotations from the *Encyclopaedia of the Social Sciences* are from J. F. Rees, "Rogers, James Edwin Thorold (1823–90)," vol. 13, p. 417, and Willard L. Thorp and George R. Taylor, "Prices," vol. 12, p. 377.

THE CITY BLOCK AS A UNIT FOR RECORDING AND ANALYZING URBAN DATA

EDWARD B. OLDS, *Research Director*
Social Planning Council, St. Louis, Missouri

Tabulations by city blocks make possible many uses of small area data beyond those which can be made from census tract tabulations. Block data can be economically analyzed and summarized by the use of summary punched cards. Some uses of block data are illustrated from the St. Louis experience. Suggestions are presented for new census data needed by blocks. The publication of a local block map and street directory facilitates the compilation of new data to supplement those obtained from the decennial census. Despite the many uses and advantages of block data, they do not replace census tract tabulations which meet a somewhat different need.

THE NEED FOR some type of standard unit geographic area for recording data about a city is apparent to many research workers and administrators faced with the problem of drawing conclusions from statistics about the city. The census tract, popularized by Dr. Walter Laidlaw and later by Howard Whipple Green, has served as the most generally accepted statistical unit area for American cities. It was developed as a compromise device to facilitate the analysis of population trends and characteristics in sections of cities. Honest attempts have been made to define census tract boundaries so that they include territory with reasonably homogeneous characteristics and with a population of from 3,000 to 6,000 persons. Unfortunately, characteristics change and what were once good boundary lines in terms of economic or ethnic indicators are not always good boundary lines. To preserve comparability from one census to another, it is necessary to keep census tract boundaries intact except for changes in city limits. However, within limits, the census tract served to reveal gross average differences between major sections of cities, as well as trends from one census to another. The census tract has admittedly an important place in the analysis of urban data, since it is small enough to show up differences between major sections of cities, and yet large enough to be easily manipulated without considerable expense. In large cities, such as New York and Chicago, it has been found necessary to combine census tracts into statistical or community areas providing more adequate bases for the computation of death rates and simplifying the mechanics of presenting and interpreting data about sections of the city.

Just as the magnifying glass is not completely replaced in usefulness by the microscope, so the usefulness of the census tract is not displaced by a more minute and precise unit area. But there are many uses of spatially ordered data about the city which can be made when they are available by a smaller unit area than the census tract. If it is conceded that a smaller unit area is desirable, the question comes up as to how much smaller the area should be and how it should be defined. Should it be something like a precinct used in organizing elections, a "beat" in police circles, or some other arbitrary grouping of city blocks? No matter what grouping is adopted, there will always be districts which cannot be made to fit the established boundaries. Does this mean that the problem is insoluble? It can be easily answered by the adoption of the city block as a unit area. Nearly all of the following types of districts are composed of city blocks.

School districts	Campaign solicitation unit areas
Water districts	Diocese and parishes
Sewer districts	Census tracts
Health districts	Neighborhood areas
Precincts and wards	Improvement districts
Police precincts and beats	Zoning districts
Meter reading districts	Fire districts
Telephone exchange districts	Library districts
Power districts	Welfare administration districts

TAX ASSESSMENT DISTRICTS

Data compiled by blocks with totals recorded in punch cards can be economically summarized by almost any of the above districts for any city. On the other hand, it is only rarely that data tabulated by census tracts or even enumeration districts can be accurately compiled according to the above types of districts. Even though an attempt is made in establishing census tracts to follow the boundaries of various administrative districts, the problem is practically insoluble without shifting the boundaries of the districts.

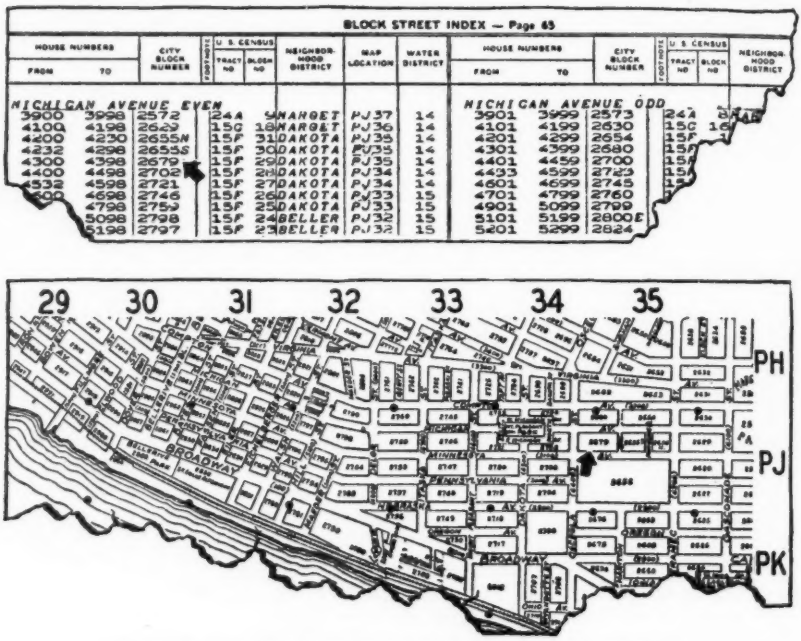
The proponents of census tracts sometimes argue that if administrators will not take the trouble to change the boundaries of their districts to conform to census tract boundaries, they can do without statistical information. Such an attitude fails to take cognizance of some of the very real difficulties preventing administrative districts from being brought into congruity with the boundaries of census tracts. For example, boundaries of school districts may have to be altered as there is movement of population, to maintain the proper balance of school enrollees in each school. If one district is growing in

population while its neighbor is declining, it is obviously simpler to move the boundary of the district to correct the unbalanced enrollment situation, rather than to change the capacity of the school. Moreover, for many purposes, it is necessary to have some unit smaller than a census tract to serve as an administrative area. For example, police beats, precincts, or campaign solicitation areas must be considerably smaller than the census tract with a population usually between 3,000 and 6,000. The investment of sizeable funds in capital equipment such as telephone exchanges, power lines, sewer mains or water pipes may make it impractical to change boundaries of control areas merely to make them conform with artificial statistical areas. Unless accurate summaries can be made of the expensively compiled census information, many valuable uses are lost. Of course, in some instances it is possible to make estimates and approximations by prorating census tract data or by using overlay maps. But as business and government become more scientific, there is increasing demand for *accurate* information on which to base future plans and policies. By use of the block summary punch cards, accurate summaries can be obtained economically without excessive cost beyond the cost of coding the original data in terms of blocks. In relation to the total cost of training enumerators. conducting the canvass, designating areas, coding and tabulating, the preparation of block summary punch cards is not excessive. If a five or ten per cent increase in cost makes possible a many-fold increase in the uses of urban data, such additional costs should be justified.

THE ST. LOUIS BLOCK STATISTICS PROJECT

Some indication of the possibilities in the use of block statistics may be gained by examining the St. Louis experience. In the fall of 1945, the local committee on census enumeration areas called the Metropolitan St. Louis Census Committee, obtained the cooperation of 21 business, government, welfare and educational establishments in sponsoring a local block statistics project. This involved purchasing a deck of block summary cards for St. Louis from the U. S. Census Bureau, converting the census block numbers to those used locally for over 60 years, and publishing a Block-Street Address Directory and Map. The cost of this work was largely covered by the sale of directories, maps, and sustaining memberships. The form of the directory and map, which was published by the offset reproduction of machine listings, is indicated in Figure 1. The map location code facilitates locating a specific block on the block map. It is also used to sort and list cards in geographic sequence to improve the efficiency of

mapping. The neighborhood district name symbol and water district code were included in the directory to satisfy two agencies which assisted considerably in its compilation, the City Plan Commission and the St. Louis City Water Department.



BLOCK STREET INDEX — Page 65

HOUSE NUMBERS FROM	HOUSE NUMBERS TO	CITY BLOCK NUMBER	FOOTNOTE	U. S. CENSUS TRACT NO	U. S. CENSUS BLOCK NO	NEIGHBORHOOD DISTRICT	MAP LOCATION	WATER DISTRICT
MICHIGAN AVENUE EVEN								
3900	3998	2572		244		SHAHGET	PJ37	14
4100	4198	2629		15C	16	WARGET	PJ36	14
4200	4230	2655N		15F	31	DAKOTA	PJ38	14
4232	4298	2655S		15F	30	DAKOTA	PJ35	14
4300	4398	2679		15F	29	DAKOTA	PJ35	14
4400	4498	2702		15F	28	DAKOTA	PJ34	14
4532	4598	2721		15F	27	DAKOTA	PJ34	14
4600	4698	2746		15F	26	DAKOTA	PJ33	15
4798	4798	2759		15F	25	DAKOTA	PJ33	15
5098	5098	2798		15F	24	BELLER	PJ32	15
5198	5198	2797		15F	23	BELLER	PJ32	15

HOUSE NUMBERS FROM	HOUSE NUMBERS TO	CITY BLOCK NUMBER	FOOTNOTE	U. S. CENSUS TRACT NO	U. S. CENSUS BLOCK NO	NEIGHBORHOOD DISTRICT
MICHIGAN AVENUE ODD						
3901	3999	2573		244	8	
4101	4199	2630		15C	16	
4201	4299	2654		15F		
4301	4399	2680		15F		
4401	4459	2700		15F		
4493	4599	2723		15F		
4601	4699	2745		15F		
4701	4799	2760				
4901	5099	2799				
5101	5199	2800E				
5201	5299	2824				

FIGURE 1. ILLUSTRATION OF FORM OF BLOCK—STREET DIRECTORY AND MAP

The use of locally established block numbers facilitates obtaining and compiling current local information. The chief data compiled regularly are:

1. Number of dwelling units in new homes for which building permits have been issued.
2. Number of dwelling units in homes for which demolition permits have been issued.
3. Number of white and Negro children enrolled in public elementary schools.

These data are useful locally to provide some indication of marked increases or decreases in population in particular neighborhoods. The school enrollment data are particularly useful in revealing annual shifts in the location of the Negro population of St. Louis. Since the

city and school authorities routinely code their records by city block numbers, the cost of making block tabulations using punch card machines is comparatively small. Summaries are tabulated by census tracts, neighborhood districts, census districts, precincts, and wards. The preparation of these summaries is facilitated by the use of a master deck of cards containing a series of code punchings signifying to which census tract, neighborhood district, etc., the particular block belongs.

For recording summary statistics about blocks, use is made of a specially printed card illustrated in Figure 2. The fields lettered from A to Q are used for quantities, such as the number of dwelling units, white school children, or dwelling units constructed in 1944. The fields printed with city block number, U. S. census tract number, block number, etc. are used for standard area codes. Reproduced decks of cards in the master file can be prepared for use by members having their own machine facilities.

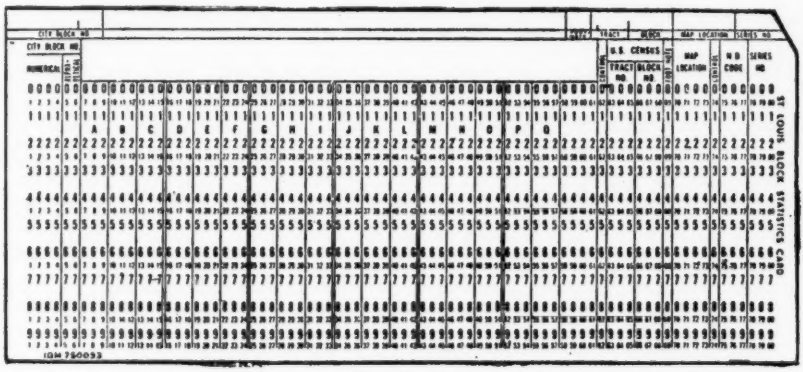


FIGURE 2. PUNCH CARD FORM USED FOR RECORDING VARIOUS TYPES OF STATISTICAL DATA FOR CITY BLOCKS

ECONOMIC RATING OF BLOCKS

To provide a convenient means of classifying addresses by economic status, a block economic rating on the basis of 1940 rents was prepared. The block summary cards from the 1940 census contained information on the average rent in each block. These cards were sorted by this average rent, listed, and at the same time the number of dwelling units in each block was cumulated. Those blocks with the lowest rents which included one per cent of the dwelling units in the city, were given a code of "01." The blocks with slightly higher rents which included another one per cent of the homes, were given a code of "02." This process

was continued until the blocks with the highest rents, which included one per cent of the homes, were given a code of "100." Addresses coded in terms of this block economic index, can be conveniently grouped into economic tenths, fifths, thirds, etc. Figure 3 illustrates the relationship between the block economic code and average rents.

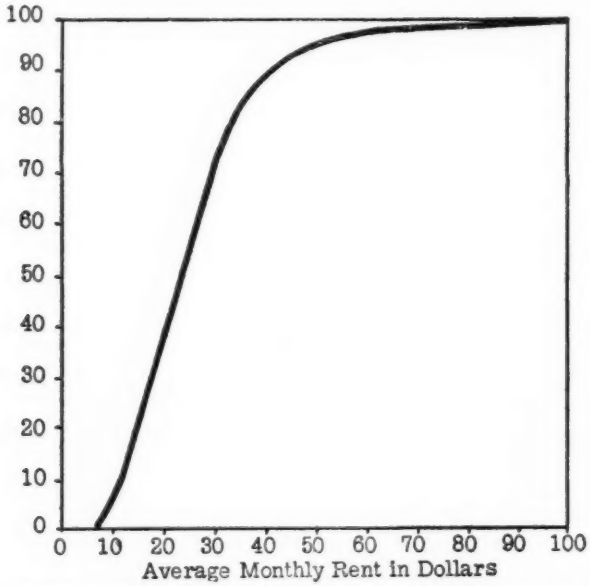


FIGURE 3. CUMULATIVE PER CENT OF HOMES IN ST. LOUIS BLOCKS WITH LESS THAN SPECIFIED AVERAGE RENTS 1940

Some of the uses which have been made of the block economic code may be of interest. In a public opinion survey, a random area sample was selected, using blocks as primary sampling units. Since it was not practical to make follow-up calls on every family in the sample, the possible biasing effect of differences in the percentage responding from low and high income areas was controlled by means of the block economic code. Tabulations were made of the number of families in the sample from five groups of economic areas classified by means of the block economic code. The distribution of usable questionnaires from these five groups of economic areas was also determined. Any significant differences in the distribution of questionnaires and the distribution of the families in the sample were corrected by obtaining more interviews. In this way, the economic composition of the families

represented by the questionnaires analyzed was kept close to the composition of the population.

In a study of subscriptions to the Community Chest obtained through neighborhood solicitation, the block economic code was used to provide an index of economic status for each solicitation area. A comparison between this economic index and subscriptions per family, indicated a marked association as illustrated in Figure 4. This information was helpful in determining the areas where neighborhood solicitation produced insufficient returns to justify the costs involved.

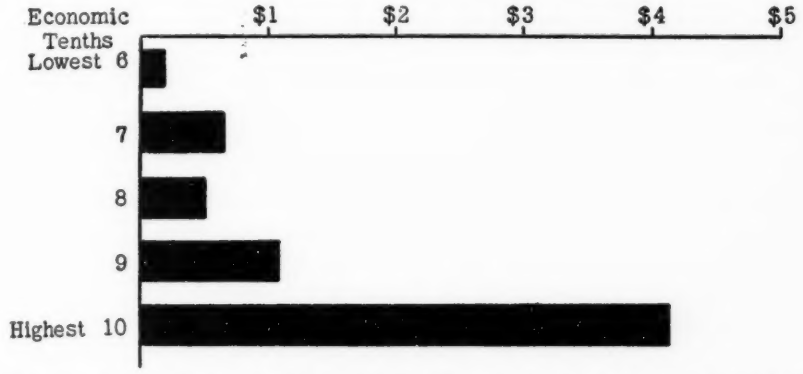


FIGURE 4. AVERAGE 1947 COMMUNITY CHEST GIFT PER FAMILY IN NEIGHBORHOOD SOLICITATION OF PART OF ST. LOUIS CLASSIFIED BY ECONOMIC TENTHS

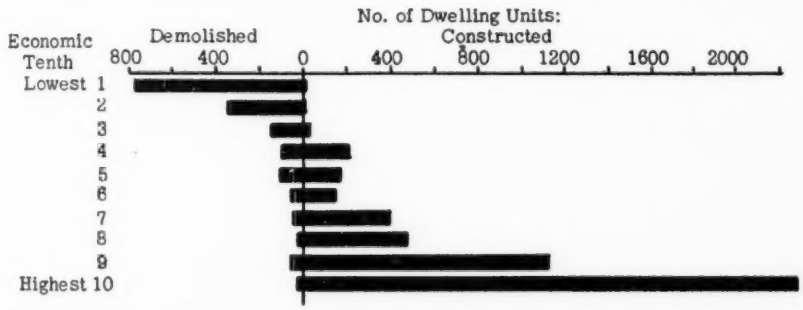


FIGURE 5. DWELLING UNITS CONSTRUCTED AND/OR DEMOLISHED UNDER PRIVATE AUSPICES IN ST. LOUIS ECONOMIC TENTHS DURING THE PERIOD 1940–46

In a study of dwelling units constructed during the period 1940 to 1946, the economic status of the blocks in which the construction took place was determined by the block economic code. A tabulation of the

units constructed and demolished according to economic tenths (each tenth containing one-tenth of the 1940 dwelling units) indicated a highly skewed distribution as shown in Figure 5. The areas with the

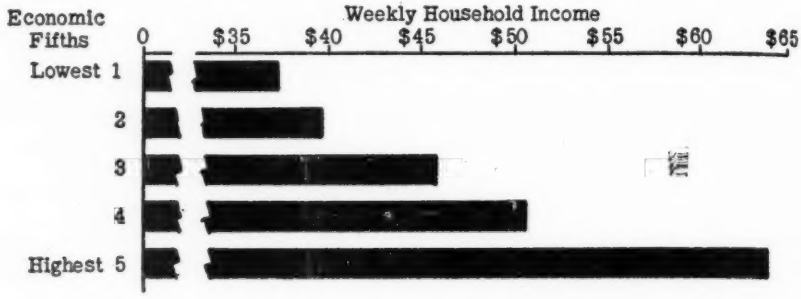


FIGURE 6. MEDIAN WEEKLY HOUSEHOLD INCOME REPORTED IN JULY 1947 BY FAMILIES IN EACH ECONOMIC FIFTH OF ST. LOUIS

highest economic status had the most building, and the areas with the lowest economic status had the least building going on. On the other hand, the high economic areas had the least demolition of homes while the low economic areas had the most demolition taking place.

Several applications of the block economic code have been made in classifying addresses in St. Louis City according to economic status. Although considerable error can result through such a method for estimating the income status of an individual, it is believed that for groups of persons or families, this method provides a fairly reliable index. Figure 6 illustrates the relationship that was found between average income reported by families in a sample survey and the block economic code based on 1940 rents. The interviews with the families were conducted in July 1947, and each family was asked to indicate in which one of the following classes their household income would fall:

Under $25 per week
$25 to $49 per week
$50 to $99 per week
$100 or more per week

A study of the economic status of the blocks in which more than 40 per cent of the population was Negro, indicated a heavy concentration of Negro blocks in the lower economic brackets. As shown in Figure 7, none of the Negro blocks were in the highest economic tenth, while 20.3 per cent of the Negro homes were in the lowest economic tenth.

Although no analysis of vital statistics data has been made, using the economic code, it is believed that highly significant differences

would be found from a comparison of life expectancy, infant deaths, etc., in low and high income areas. Such analyses require the coding of

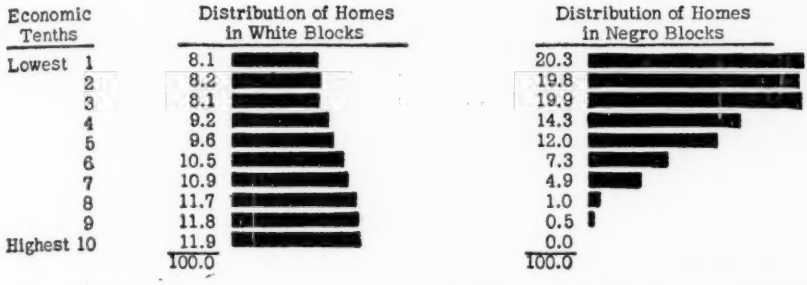


FIGURE 7. COMPARISON OF DISTRIBUTION OF HOMES IN WHITE AND NEGRO AREAS OF ST. LOUIS ACCORDING TO ECONOMIC TENTHS

Note: Each economic tenth included 10 per cent of the homes in St. Louis in 1940.

births and deaths by blocks, as well as the tabulation of population census data by blocks.

COMPILATIONS FOR SPECIAL AREAS

The block data on punch cards have been used to obtain summary tabulations of housing, school enrollment, and building permit statistics for such areas as neighborhood districts, precincts, and wards. The City Plan Commission has divided the city into 99 neighborhood and industrial districts, basing the determination of boundaries upon such factors as major streets, railroads, proximity to parks and playgrounds, land use, etc. Until block statistics were available, it was not possible to obtain accurate housing statistics for these basic planning units. As part of the St. Louis block statistics project, summary tabulations of the 1940 housing census block statistics were obtained. These included the following data:

Residential structures	Negro families
Dwelling units (homes)	Homes with more than 1.5 persons per room
Owner occupied homes	Homes needing major repairs
Tenant occupied homes	Homes without private bath
Homes built 1930 to 1939	Average monthly rent of homes
Homes built 1920 to 1929	Total rent
Homes built 1900 to 1919	Number reporting rent
Homes built before 1900	

VOTING BEHAVIOR STUDY

Summaries of housing data have been made for the precincts and wards of St. Louis. The summarized data were then used to compute

octile ratings for each precinct in each of four housing factors. This was done by computing percentages for each precinct, ranking the percentages, and then grouping the ranked precincts into eight groups. The four factors were as follows:

1. Per cent of homes owner occupied.
2. Per cent of homes built before 1900.
3. Per cent of public school enrollees who were Negro in Nov. 1946.
4. Average rents.

The block data were also used to prepare estimates of the population 21 and over in each precinct as of January 1, 1948. These estimates were based upon the current estimated number of families, using the 1940 count of dwelling units, plus units represented in building permits issued since 1940, and less dwelling units represented in permits for demolitions since 1940. The estimated number of families was multiplied by the 1940 ratio of population 21 and over, to families in the nearest census tract. The sum of these products for the city was compared with the estimated population 21 and over in the city. The provisional estimate for each precinct was then multiplied by a correction factor so that the figures finally used add up to the estimated population 21 and over in the city. While this method is subject to considerable error, it was considered more reliable than any other available method for obtaining a current estimate of population 21 and over. Percentages and octile ratings were then computed for the proportion of the voting population registered to vote. A series of 28 other octile ratings was computed from the election statistics on civic issues, as well as for political parties in eight elections held since Nov. 1944. Comparisons were not made prior to this time because of non-comparable precinct boundaries. The data prior to 1948 for each precinct were summarized on one specially printed tabulating card. Other punched cards were used to list the statistical data onto the printed card. Complete sets of 784 precinct data cards were turned over to the sponsors of the project. Other sets can be prepared economically from the master cards. The top line of each card contains a series of 23 octile ratings. The percentages upon which these ratings were based are specified by small numbers printed below each rating and in the lower left corner of each percentage cell. A set of punched cards containing the octile ratings was used for an intercorrelation analysis, using the tetrachoric correlation method. This analysis indicated the following significant relationships between the housing and voting indexes:

1. Democratic precincts tended to remain Democratic and Republican precincts tended to remain Republican.
2. Precincts with a large proportion of the population registered tended to have a large proportion of the registrants voting in each election.
3. Areas with high home ownership had a larger proportion of the population registered than areas with low home ownership.
4. High rent areas were more inclined to vote Republican than low rent areas.
5. Areas with many old homes opposed daylight saving time and a new state constitution.

TABULATION OF SCHOOL ENROLLEES

Each year in November, the Board of Education asks each elementary school to prepare a report listing the block numbers in which their pupils reside and the number of pupils in each block. The Block-Street Address Directory is used in coding addresses by blocks. Since St. Louis has a completely segregated school system, it is possible to make tabulations of these data to show the number of white and Negro school enrollees in each block. Such tabulations have been made for each of the following years:—1941, 1945, 1946, 9147. From these data by block, it has been possible to prepare a map which shows the trend of movement of the Negro areas in St. Louis during the period 1941 to 1947. Figure 8 illustrates the *Post Dispatch* map[1] drawn from the more precise block map published in two colors by the Social Planning Council and the Urban League. The housing statistics and land use data by blocks were used to compute differences in dwelling units per residential area between white and Negro areas.

ESTABLISHING CAMPAIGN DISTRICTS

Uses of the Block-Street Address Directory can be shown by describing a project involving the grouping of addresses of campaign prospects into convenient solicitation control areas. Cards were punched alphabetically giving the names and addresses of about 20,000 prospects. The punching of names and addresses was needed for the preparation of prospect lists, pledge cards, and mailing strips. In punching the addresses, house numbers were punched in one field while street names were punched in another field. Accordingly, it was possible to mechanically sort the cards by street and house number, keeping the

[1] The *St. Louis Post Dispatch* carried a feature article by Richard G. Baumhoff on this study in the Sunday issue, Aug. 15, 1948.

odd house numbers separate from the even house numbers. The use of the Block-Street Address Directory, together with a listing of these sorted address cards, privided a highly efficient means of establishing ing the block codes. In marking the list with the block code, it was found that usually from five to 20 adjacent listings would be in the same block. The punching of the block code into the cards was accomplished by manually filing the cards behind pre-punched block

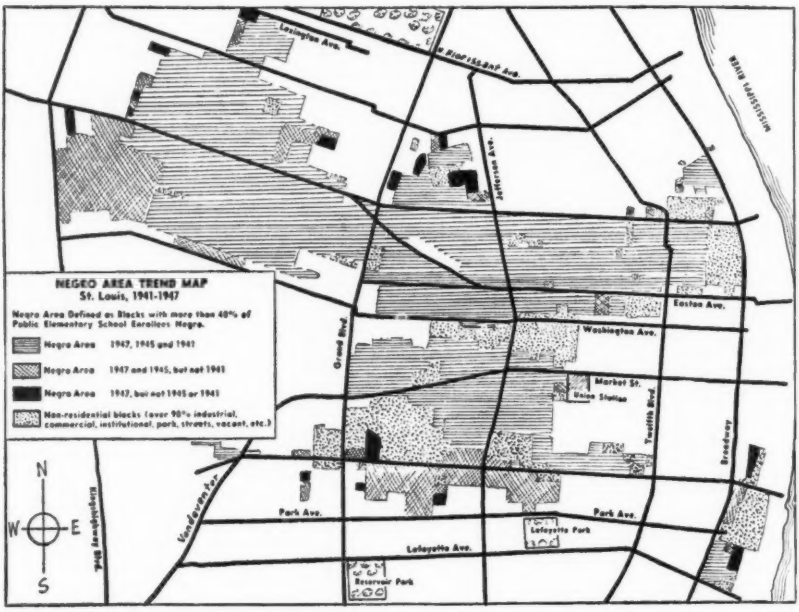


FIGURE 8. NEGRO AREA TREND MAP

The author wishes to acknowledge with thanks the permission granted by the *St. Louis Post Dispatch* to publish this map.

master cards and then intersperse gang punching the cards. The cards were then sorted down and tabulated by block. Work maps were posted with the number of prospects in each block and area boundaries were drawn to include the proper number of prospects in each area.

MAKING SPOT MAPS

In St. Louis, the city block numbers consist of four numerical digits and two alphabetical suffixes. Although this makes a rather cumbersome number, it is useful because many city records and maps are referenced with these official city block numbers. However, to locate

any given block efficiently, it is necessary to have what is called a supplementary "map location" number. This number consists of two letters followed by two numbers, like "PQ42" which defines a particular square of land in the city with sides one-fourth mile long. Every block is assigned to one specific square on the basis of where the majority of its area is located. Cards punched with city block number or census tract and block number can be automatically gang punched with this map location number and other area codes at one operation. They can then be sorted and listed in geographic columns and rows which greatly facilitates the spotting of block maps to show the accurate distribution of addresses.

POTENTIAL USES OF BLOCK DATA

The foregoing examples represent only a few of the possible uses of population and housing statistics by blocks. A consideration of these uses should suggest many others which would be made if data were available uniformly for every metropolitan area including the suburbs as well as the central city. The fact that accurate summaries can be made economically for a wide variety of administrative and study areas opens up uses which cannot be made of census tract statistics.

Some of the uses which could be made through a more general availability of block statistics are indicated in the following list:

Determination of fire, theft, and life insurance risks in different types of neighborhoods.
Studies of land values as related to population, sales, etc.
Determination of business areas of the Metropolitan District.
Planning changes in the location of transportation and utility lines.
Appraisal of property for loans or taxation.
Determination of cost of governmental and philanthropic services in each section of the city as compared to tax income obtained.
Planning optimum location of public, private, or commercial facilities for recreation, education, sales, health or welfare service, etc.
Estimates of sales or consumption using block statistics in the design of the sampling plan.
Indexing detail real estate or land use maps.
Determination of optimum districts for neighborhood improvement, police beats, meter readers, relief investigators.

SUGGESTED NEW CENSUS BLOCK DATA

One of the limitations upon the use of block statistics is the paucity of information generally available by blocks. It is believed that there could be a considerable increase in the variety of data tabulated from the decennial census without greatly increasing costs. If territory is assigned enumerators by blocks (as was done in the 1940 housing

census) the punching of a block designation in all punch cards made from the schedules would be comparatively simple. Then special tabulations could be run by blocks or groupings of blocks for any desired detail. Routinely, it is believed that certain summary information could be tabulated by blocks recording the totals in summary cards. The following items could be recorded in three decks of such cards:

Housing data (in order of importance):
- Number of dwelling units.
- Contract or estimated monthly rent.
- Owner occupied dwelling units.
- Dwelling units occupied by non-white persons.
- Dwelling units built before 1900.
- Dwelling units with no private bath.
- Number of dwelling structures.
- Dwelling units according to type of structure (3 groups).
- Dwelling units without private flush toilet.
- Dwelling units without running water.
- Dwelling units without mechanical refrigeration.
- Dwelling units without central heating.

Population data (in order of importance):
- Number of persons.
- Age distribution (5 groups)
- Sex and color (4 groups).
- Population 25 and over by years of school completed (6 groups).
- Population 14 and over by employment status and sex (12 groups).
- Employed persons 14 and over by major occupation group (9 groups).

These data could be economically published by listing them on plastic offset plates and reproducing several hundred copies for sale to users at a charge set to write off the publication cost. Users would be encouraged to purchase decks of cards on printed card forms clearly indicating the information punched in the cards. Users could also be supplied at cost with block maps of adequate scale for work purposes, reproduced through a blue print process from masters kept in the Census Bureau. Such maps should be in sections that would be small enough to be handled easily on normal size desks and drafting tables, and yet so made that they could be easily assembled to make up a one-piece map for a Metropolitan District. A certain amount of skilled consultation service should be made available by the Census Bureau to help users in making the best possible use of the block data.

SUGGESTED LOCAL BLOCK TABULATIONS

The preparation of block statistics as outlined above would require

local committees in each community to help in making the best uses, as well as in promoting the compilation of local material. One of the first projects for each city committee would be the purchase of decks of the summary cards and sets of block maps in the form of negative blue print masters. Another project would involve the compilation of a local block-street address directory to facilitate the compilation of local data by blocks. The promotion of local tabulations should include consideration for obtaining the following types of information:

Building erections and demolitions.
School census or school enrollment data.
Land use statistics.
Police arrests.
Juvenile delinquency cases.
Births and deaths.
Persons receiving welfare services (chronically ill, tubercular, mentally ill, general hospitalization, foster home placement, etc.).
Old age assistance, aid to dependent children, and general relief cases.
Tax assessments and collections.
Fire losses.

The financing of local projects can be handled in various ways, depending upon the community. Generally, it is possible for each administrative agency to include in its budget, small amounts sufficient to cover the processing of statistics, falling within its jurisdiction. Sales of directories can be used to cover the cost of their compilation and publication. Contributions from utilities, banks, chambers of commerce, universities, foundations, real estate firms, etc., can be used to write off the cost of local projects. However, it is necessary to have interested and competent leadership for local committees. Such persons may be found in a local city plan commission, council of social agencies, university, chamber of commerce, utility, board of education, etc. If interest warrants, it may be possible in certain communities to establish agencies equipped with staff and machinery for the most efficient processing of statistical data having general community-wide significance. Block statistics would be one of the kingpins in such a community research agency.

National agencies and concerns should find considerable uses for block statistics when they become uniformly available together with adequate maps and street address directories. Survey and polling organizations should be able to effect economies and improvements in their work through the use of block statistics. With adequate materials and interpretation, it should be possible to cover some of the added

census cost of block statistics through the sale of cards, maps, and listings.

COMPARATIVE ADVANTAGES OF CENSUS TRACT AND BLOCK DATA

From the St. Louis experience, we find that block statistics are essential for the following types of analyses:

1. Compilation of census statistics for administrative areas which are not multiples of census tracts or enumeration districts.
2. Compilation of census statistics for areas within a specified distance or travel time of a particular geographic location (half mile or mile).
3. Classification of addresses according to economic status based upon economic index computed for blocks.
4. Appraisal of neighborhood characteristics in immediate vicinity of a specified address.
5. Determination of exact boundaries of areas inhabited predominantly by a particular ethnic group.
6. Design and selection of area samples for use in factual, attitude, and opinion surveys.

Census tract data are inadequate, although better than no data, for analyses such as the above. However, census tract data are preferable to block data for analyses such as the following:

1. Community studies of districts or sections within an urban area when the population of the district is over 20,000.
2. Computation of ratios such as tuberculosis death rates for sections of a city.
3. Presentation of a general view of the variation from community to community within a city in significant population or housing characteristics.

In large cities even census tracts are too small for use in analyses such as the above.

In conclusion, statistical tabulations of selected types of data by city blocks supplement, rather than displace tabulations by census tracts. Wherever possible, administrative districts within a metropolitan area should be established as multiples of census tracts. Block statistics should be used only for analyses which require greater geographic detail than can be provided by census tract data. The judicious use of block and census tract statistics can make a noteworthy contribution to the more scientific administration of business and governmental services.

THE RELATION OF THE NET REPRODUCTION RATE TO OTHER FERTILITY MEASURES

T. J. WOOFTER

Recent population literature has been critical of the net reproduction rate on the grounds that it is based only on the female population, that it assumes the invariable continuation of the reproductive situation of a single year, and that it ignores the past childbearing experience of the generation.

Alternative measures are:

Male reproduction rates;

Marital reproduction rates, which are of two types—(a) those showing the birth rates of a single year standardized for duration of marriage, and (b) those showing the number of children ever born to women who have been married for varying periods of time.

Generation rates, which are based on the total number of children ever born to a generation of women who have completed the childbearing period.

Standardized quota reproduction rates are proposed in order to preserve the generation principle, but center the experience measured closer to the current year.

Rates adjusted for the order of birth of children, as proposed by Whelpton, may be calculated either from the base of a single calendar year or from the complete experience of a generation.

All of these rates may be classified into two types: (1) Those which depend on the birth rates of a single year, and (2) those which cumulate the experience of a group over a period of years. The former are more sensitive to short-time changes in the birth rate, while the latter provide a longer and more stable base for measuring trends.

I. CRITICISMS OF THE NET REPRODUCTION RATE[1]

SINCE THE FIRST exposition of the net reproduction rate and Lotka's demonstration of its relationship to the true rate of natural increase, it has been *the* measure of fertility used by demographers and has provided the techniques for some of the most penetrating analyses of fertility which have been produced in the two past decades. Recently, however, the dissatisfaction with the net reproduction rate as

[1] The most recent discussion of the net reproduction rate, which has been presented by Lotka, was contained in a paper soon to be published in the Proceedings of the Meeting of the International Statistical Institute, held in Washington in 1947. Cf. also *Journal of the American Statistical Association*, June 1936, "The Geographic Distribution of Intrinsic Natural Increase in the United States and the Examination of the Relation of Measures of Net Reproduction."

the measure of fertility, especially of fertility trends, has led to a reappraisal of the possibilities of alternative measures.

The gross reproduction rate is the sum of the age-specific female fertility rates of all women. The net rate is derived by multiplying the age-specific female fertility rates by the corresponding female survival rates and summing the products; thereby yielding a ratio of the number of daughters who will survive to the age at which their mothers gave them birth to the number of women in that age, the underlying assumption being that fertility rates or the mortality rates of a particular year will remain constant.

Some of the criticisms of the net reproduction rate seem to have been caused by a confusion as to the uses to which the rate may be put. It has, in other words, sometimes been carelessly used by those who are not familiar with all of its assumptions. This has been especially true when attempts have been made to use the customary reproduction rates for prediction. Three general uses to which such rates have been put may be distinguished:

First, they are used for the simple comparisons of two or more populations by means of rates relating to the same year, or of short-term variations within the same population. Examples are the comparison of urban and rural fertility in 1940, or the comparison of fertility in a particular area in two successive years. Obviously, the accomplishment of this purpose often requires a measure of fertility performance at a particular time which is sensitive to all of the combined factors affecting fertility. For such general purposes, the crude birth rate is often satisfactory, or the age-standardized female birth rate, which is the gross reproduction rate, may be used.

The second objective of such rates is the analysis of the factors affecting fertility, by attempting to isolate the effect of the principal factors being studied. It is especially important, when correlations of economic or social variables are made with the birth rate, that as many other factors be standardized as possible. The requirement for these purposes is a series of measures standardized for different factors in order that the difference between standardized and unstandardized rates may reveal the fluctuations which arise from the factor under consideration.

The third purpose is the endeavor to predict the future by the net reproduction rate of a particular year or of a series of years. This interpretation of the rate sometimes creeps into the thinking of the layman and even of the uninitiated investigator. This confusion is understandable, since the very term "reproduction" implies a look

into the future. The question is often asked in this way: "Is this population reproducing itself?" This is only another way of asking the question, "Will the next generation be as large as this generation?" In interpreting such a dynamic situation, some demographers have used a measure which assumes static conditions.

All through the recent European literature which criticizes the net reproduction rate we find such terms as "invariable fertility," or "relation of the tendency value to the situation of the moment," or "the long-term prospects of population growth"; the implication being that, inasmuch as the net reproduction rate does not satisfactorily provide such prediction factors, there should be a rate which does. This concept, obviously, gets into the realm of philosophical reasoning. Are there in various populations inherent fertility trends which change slowly, but persist for a long period of time? In recent discussions of fertility measures, one gets the impression that some demographers think that there are such invariable fertility trends. If we assume that there is some underlying fertility pattern, there is difficulty not only in choosing the proper measure to characterize it, but also, if prediction is attempted, the investigator is confronted with all the technical difficulties which beset extrapolation in any field.

The three principal objections which have been raised against the net reproduction rate are: (a) It applies only to females and takes no account of the sex ratio and difference in ages of fathers and mothers; (b) that, for this rate to be meaningful in practical terms, it assumes the invariable extension of the reproductive situation of a particular moment; and (c) that it ignores the past in that it does not make allowance for the influence of the past fertility experiences of the women who have children in a particular year.

Even though the recent reversal of the drop in the birth rate may be temporary, it has been sharp enough to cause a new crop of analyses of fertility and to intensify experimentation with various measures which might explain the phenomena. In the case of this country, the native white gross reproduction rate fell from about 160 in 1915 to 104 in 1936 and rebounded to 159 in 1947. This violent fluctuation is in itself concrete evidence of the small probability that the rates of any one year will remain invariable, which is a basic aassumption in the use of the net reproduction rate as a measure of generation reproduction.

Critics of the net reproduction rate argue that, while long-time trends may continue with some stability (if such long-time trends can be discerned), still the variations in economic and social conditions

and in familial attitudes change rapidly and have different effects on short-term fluctuations in fertility. Hence, they have endeavored to analyze the effect of these short-term variables and, secondarily, have groped for a measure of what might be characterized as the underlying fertility trend. The lines of approach which are treated in the following pages are: (1) Male reproduction rates; (2) nuptial reproduction rates; (3) generation reproduction rates; (4) cohort replacement rates; and (5) rates adjusted for order of birth of children.

It is not possible in the scope of one article to develop the details of each of these types of measures, but their chief characteristics may be described briefly, especially their relationship to the net reproduction rate and the availability of data for their calculation.

II. MALE NET REPRODUCTION RATES[2]

In theory, as well as in practice, the male net reproduction rate is similar to the female, except that it measures the number of sons born to 100 fathers who will survive to the age which their father had attained when they were born. It would appear that, if the population is assumed to be tending toward an equilibrium, when 100 women produce 100 surviving daughters, then it would also be tending toward an equilibrium when 100 fathers produce 100 surviving sons. If, therefore, for a particular year (as in England in 1938) the paternal net reproduction rate was .881 and the maternal only .808, which is the most pertinent as the basis for judging the effect of 1938 fertility conditions on future trends?

Difference between the mean length of male and female generations is not the only cause of discrepancies. There is a difference in the proportion married at various ages; there is a difference between the distribution within the childbearing ages of men and women; and there are differential mortality rates. All of these factors are reflected in differences between male and female fertility rates. While there is no apparent reason for choosing the rate of one sex as more useful than the rate of another, comparisons between the two are revealing as the effect of the sex ratio and differential age of marriages in the population.

---

[2] Myers, R. J., "The Validity and Significance of Male Net Reproduction Rates," *Journal of the American Statistical Association*, Vol. 36, No. 214, June 1941.

Hajnal, J., "Aspects of Recent Marriage Trends in England and Wales," *Population Studies*, Vol. 1, No. 1, June 1947.

Tietze, Christopher, "Differential Reproduction in the United States," *American Journal of Sociology*, Vol. 49, No. 3, 1943.

———, "Differential Reproduction," *Milbank Memorial Quarterly*, Vol. 19, No. 3, July 1939.

Karmel, P. H., "The Relations Between the Male and the Female Reproduction Rates," *Population Studies*, Vol. 1, No. 3, Dec. 1947.

III. NUPTIAL REPRODUCTION RATES[3]

Nuptial reproduction rates have been in use for some time, but, in recent years, experimentation with these rates by European scholars has increased to a great extent. In fact, some European demographers assert that such rates are the most satisfactory which have been evolved to date. As a result, their techniques have been considerably refined, the chief modification being a refinement to allow for duration of marriage. The previously used nuptial rates (unadjusted for duration) were constructed on the same theoretical framework as a conventional net reproduction rate, in that they were derived from age and marital specific rates based upon the fertility and mortality experience of a single year. However, they had the advantage of being specific for nuptiality and allowing the student to isolate this factor for special study. The use of such a rate as a basis for prediction is analogous to the use of the net reproduction rate, in that it is assumed to indicate the eventual rate of increase of a population with stable age structure and invariable extension of the fertility, marriage, and mortality rates of the year in question. For practical purposes, this is evidently open to the same objection which the advocates of nuptial rates have made against the conventional net reproduction rate.

Refinement of the general nuptial rate to allow for duration of marriage is made in two ways:

(1) Fertility rates according to nuptiality are calculated for each age and duration of marriage, and combined into a rate for all married women by the use of a nuptiality table. These rates are again based on a single calendar year's experience and, therefore, open to the same objection which we continually emphasize; namely, that, as a predictor, this assumes the invariable extension into the future of a single year's fertility, mortality, and marital rates, with the added objection that another factor (fertility rates by duration of marriage) is also assumed to remain constant. (For illustration, cf. Table 1.)

An ingenious refinements of such rate has recently been proposed by Hyrenius.[3] This method takes into account both the male and female age and marital distribution by duration of marriage. Its derivation is described by him as follows: "(a) The elaboration of an index of the proportion of sexes among the non-married persons within

---

[3] Glass, D. V., *Population Policies and Movements in Europe*, Appendix, pages 399–405.

Hajnal, J., "Analysis of Recent International Recovery in the Birth Rate," *Population Studies*, Vol. 1, No. 2, September 1947.

Quensel, Carl-Erik, "Population Movements in Sweden in Recent Years," *Population Studies*, Vol. 1, No. 1, June 1947.

Hyrenius, Hannes, "La Measure de la Reproduction et de Accroisment Naturel," *Population*, April–June 1948.

certain age limits; (b) the calculation on the basis of this index and of observed nuptiality rates of a table adjusted for feminine nuptiality and the distribution of the newly married by age; (c) construction, on the basis of mortality and divorce statistics, of a table of attrition of marriages for various groups of ages at marriage; adjustment of an

TABLE 1*

FEMALE MARITAL FERTILITY RATES BY DURATION OF MARRIAGE
(Age of wife: 15–49 years)

Duration of Marriage (single years)	Fertility Rate Per 1,000 Married Women			
	1933	1939	1941	1943
0–	363	320	325	376
1–	197	198	212	239
2–	167	171	168	197
3–	144	154	151	195
4–	129	133	133	165
5–	112	123	115	150
6–	100	109	101	135
7–	89	95	90	123
8–	82	83	80	104
9–	73	74	70	90
10–	67	66	62	82
11–	59	58	56	72
12–	55	54	51	63
13–	52	48	45	54
14–	48	45	39	46
15–	40	34	35	39
16–	38	31	32	34
17–	32	30	27	30
18–	25	25	22	24
19–	26	21	19	21

* From "Population Movements in Sweden in Recent Years," by Carl-Erik Quensel, *Population Studies*, June 1947, p. 34.

analytical function to this table; (d) calculation, by combining the distribution by age of the newly married with the table of attrition of marriages, of the distribution of married women by age of the husband and duration of marriage; (e) calculation, on the basis of the preceding functions of the rate of reproduction and intrinsic rate of natural increase, legitimate fertility being given by age of the mother associated with duration of marriage, since illegitimate fertility is only known by age groups." This is, manifestly, a complex computation—one which requires more data than are available for most populations.

(2) When births are recorded, year by year, according to age of mother and duration of marriage, it is possible to trace a cohort of

marriages back to the date of marriage and construct a table showing the number of children ever born to women married in certain years, according to age and duration up to the most recent year available. A similar table can be constructed from census enumerations, which record data on duration of marriage, age, and number of children ever born (cf. reference above to Hajnal). (For illustration, cf. Table 2.)

TABLE 2*

AVERAGE NUMBER OF CHILDREN TO MARRIAGES, BY DURATION OF MARRIAGE, AT THE END OF THE YEARS 1933, 1939 AND 1943

(Age of wife at marriage, 20–24 years)

Duration of Marriage (years)	Average Number of Children Per Marriage at the End of		
	1933	1939	1943
1	0.45	0.39	0.43
2	0.68	0.62	0.66
3	0.88	0.84	0.82
4	1.06	1.01	1.06
5	1.24	1.14	1.15
6	1.42	1.27	1.33
7	1.58	1.41	1.43
8	1.73	1.53	1.60
9	1.89	1.64	1.62
10	2.03	1.74	1.71
11	2.17	1.85	1.80
12	2.36	1.97	1.88
13	2.53	2.09	1.95
14	2.71	2.20	2.03
15	2.82	2.31	2.06
16	2.96	2.41	2.17
17	3.10	2.52	2.27
18	3.23	2.69	2.36
19	3.40	2.83	2.45
20	3.56	2.97	2.54

* From "Population Movements in Sweden in Recent Years," by Carl-Erik Quensel, *Population Studies*, June 1947, p. 35.

Such tables have the advantage of taking into account the past fertility history of women in the cohort by duration of marriage. Calculations based on such tables broaden the base period from the fertility experiences of the most recently recorded year to that of the whole range of fertility history of married women in the childbearing ages. By making this shift in emphasis, such rates have a distinct advantage in that they predict the extension of a much longer and more stable experience than that of a particular year. Also, they predict the extension of an experience in which the influence of previous births to the mother is taken into account. If the principal determinator of long-time reproduction is the average number of children which married

couples will rear during the whole span of their life, then the marriage rates based upon the number of children ever born by duration of marriage begin to develop a body of data which can provide the answer. However, the question of predicting fertility is broader than that of predicting married fertility alone. For general fertility, the answer must take into account variation from time to time in the proportions of married people at various ages.

Unfortunately, also, the dates at which duration of marriage-fertility data have become available are so recent, even in most European countries, that insufficient time has elapsed for a reliable trend to appear. In the United States, reliable fertility data by duration of marriage are not available in any form. Year-to-year cross sections could be obtained from Census enumerations if questions were added as to the duration of marriage and these were cross-tabulated with the items of age and number of children ever born. Addition of the question of duration of marriage to the birth certificate in this country would not by itself solve the problem, for it would have to be supplemented by accurate nuptiality tables. Otherwise, there would be no population base upon which the rate could be calculated, and technical difficulties would be injected by reason of plural marriages and broken marriages.

Besides the short span of available data, there are other technical complexities in constructing marriage-adjusted fertility rates. Not the least of these is the problem of handling illegitimate births. This difficulty is usually recognized by calculation of separate rates for legitimate and illegitimate fertility. Another difficulty arises from the dissolution of marriages. If, as indicated above, the ex-married are progressively eliminated from the nuptiality table, a bias is probably introduced, for the reason that the complete fertility of married couples who remain married is greater than that of couples whose married life is interrupted during the childbearing period by death, separation, or divorce. The Census approach offers a possibility of minimizing this difficulty by the combination of children ever born to the ex-married with children born to the married.

In the United States, the fact that Census fertility data are secured only from married and ex-married women and the fact that young children are not completely enumerated reduces the accuracy of such information from enumerations. The addition of the duration of marriage question to the birth certificate would run counter to the established social policy of eliminating facts as to illegitimacy from birth records.

Theoretically, the objection to the use of number of children ever born by duration of marriage (unless combined in some fashion with the number of illegitimate births) arises from the fact that it ignores the changes in the percentage of the population who are married, because the base of the rate is narrowed from all women to married women. The rates do, however, have the advantage of getting away from a single year's base and the additional advantages of taking into account previous fertility history and providing for analysis of the effect of duration of marriage as a separate factor.

Since post-war European investigations have relied so extensively on marriage-adjusted rates, it behooves American students to follow this development carefully to determine whether it is not desirable to develop nuptiality tables and to secure reliable data by age, birth order, and duration of marriage. Whelpton, in his recent work, has corrected hypothetical cohorts for spinsterhood, but this correction has consisted of an arbitrary reduction of 10 per cent in the married population in each age, on the ground that this is the usual proportion of those who remain unmarried until the end of the childbearing period.

IV. GENERATION REPRODUCTION RATES[4]

The assumption underlying the net reproduction rate that the fertility and mortality rates of a single calendar year will remain invariable may be obviated by accumulating the actual number of female children born in each of the years of life of a generation; i.e., the generation gross rate is the sum of the age-specific gross female fertility rates obtained at dates which are advanced one calendar year for each advance of one year in the age of the mother. Similarly, appropriate generation mortality rates may be applied to determine the number of these daughters who will survive to the age at which their mother gave them birth. Illustration of the arrangement of rates and calculation is shown in Table 3. If a reliable Census enumeration of the number of children ever born to women who have lived through the childbearing period by various dates is available, a similar but somewhat cruder measure may be calculated from the number of births reported by women above age 45. This measure corresponds fairly well to the measure calculated from actual generation reproduction frequencies, except that in the United States, the Census enumerations tend to

[4] Depoid, M., "Reproduction Nette en Europe Depuis l'Origine de l'Etat Civil," *Etudes Demographique*, No. 1, Statistique General de la France.

Woofter, T. J., "Completed Generation Reproduction Rates," *Human Biology*, Vol. 19, No. 3, September 1947.

underestimate the number of children ever born, either because of faulty memory of respondents, or because of the tendency not to report illegitimate children or children of a previous marriage.[5] Such generation measures avoid the assumption of invariable fertility and mortality rates by substituting for the rates of a single calendar year the actual age-specific fertility rates of a particular cohort of women who live from ages 15 to 45. Another advantage is that they are based

TABLE 3

CHILDREN PER 1,000 WOMEN, GENERATION BEGINNING REPRODUCTION IN 1915

Age	Calendar Years	Children Per 1,000 Women				
		1st Year	2nd Year	3rd Year	4th Year	5th Year
16–19	1915–19	58	57	56	57	52
20–24	1920–24	166	166	153	152	154
25–29	1925–29	148	142	139	131	127
30–34	1930–34	95	89	85	80	82
35–39	1935–39	51	48	46	46	45
40–44	1940–44	17	17	15	16	17

Total births per 1,000 women living to age 45	2,507
Female births per 100 women	121.6
Female survival rate to age 28, 1933 life table*	.898
Net reproduction rate	1.092

* Instead of calculating separately the survival of daughters born to mothers at each age, accurate results may be obtained by calculating the survival of all daughters up to age 28 (the average age of mother) by a survival rate appropriate to the calendar year 18 years after the generation began reproduction. (Cf. "Generation Reproduction Rates," supra.)

upon the whole universe of women without elimination of the variations caused by percentages who are married. That is to say, they measure the impact of all factors operating upon fertility during a single generation of women because they are only standardized for age and mortality. They also have the advantage of allowing for the effect of past fertility performance on the fertility rates of the moment. Consequently, the generation rate is a slowly fluctuating measure in contrast with the rapidly changing rate of the calendar-year net reproduction rate. It, likewise, allows for the changes in the order of birth of children by following one group of women all the way through the childbearing period. In fact, the author has pointed out in the article previously cited that the generation method may be applied to the birth rates of children of the first, second, third, and higher orders, as well as to the total fertility rate. In these cases, the sum of the birth

[5] For such comparisons, see "Completed Generation Reproduction Rates," cited in Footnote 4.

rates of children of all orders in each age is equal to the gross reproduction rate.

An allowance for improvement in mortality by use of a generation life table results in a marked increase in net rates over those calculated on the assumption of the invariable continuation of current survival rates. The difference may be illustrated by the effect on the women born in 1900. When these women were 15 to 19 in the years 1915 to 1919, their children had a probability of surviving to age 17½ of only .878. The children born to the same generation of women when they were age 40 to 45 have a probability of surviving to age 17½ of .948,[6] an improvement of 8 per cent in 25 years (Table 4). In fact, the children born in 1940 had a probability of surviving to age 40 which is superior to the probability that children born in 1915 would survivive to age 5.

TABLE 4

5-YEAR GENERATION SURVIVAL RATES (MA) OF WHITE FEMALE CHILDREN UP TO AGE OF MOTHERS FOR GENERATIONS BEGINNING 1915–1940 AND 1920–1940

Mother's Age When Child Was Born	Initial Reproduction Year of Generation					
	1915	1920	1925	1930	1935	1940
15–19	.878	.900	.915	.927	.938	.948
20–24	.895	.910	.922	.934	.943	
25–29	.905	.917	.929	.938		
30–34	.911	.922	.933			
35–39	.917	.925				
40–44	.916					

It has been pointed out, however, that these rates have one serious disadvantage, namely, the longer time span covered. The complete generation experience of a cohort of women can only be recorded after age 45, which means that the child bearing experience extends back for 30 years. Likewise, the children who are born to women age 45 remain at risk of death for another 45 years, making the complete span of generation replacement performance in the neighborhood of 75 years.[7] This necessitates either carrying birth rates far into the past or

[6] In making such calculations for recent generations, it is necessary to project survival rates into the future to some extent. The generation life tables used in this and the previous article by the author on this subject were calculated by using actual survival rates up to 1940 and Whelpton's medium mortality assumptions thereafter.

[7] The average span is, however, much shorter. The author has pointed out in a previous article (cf. Footnote 4) that an accurate method of calculating generation mortality of the children born to a generation of mothers is to apply the single survival rate to age 28 (the average age at which mothers bore children) from a life table applying to a date 18 years after the generation begins child-bearing; i.e., the children born to women beginning childbearing in 1915 would survive at an average rate equal to survival to age 28 in 1933. Thus survival rates do not have to be estimated after the terminal child-bearing age.

projecting them far into the future, and of making similar projections of survival rates. The whole United States was included in the birth registration area only in the 1930's. Whelpton has estimated rates back to 1920, and similar estimates for the group of mothers 15 to 19 may be made with some confidence back to 1915, with the result that only the generations beginning in 1915, 1916, 1917, 1918, and 1919 can be satisfactorily estimated. (These generations reached the age 44 in the years 1944, 1945, 1946, 1947, and 1948, and the daughters of the older mothers of these generations will be in their older childbearing ages from the years 1989 to 1993; but the average mortality of daughters of these generations may be closely approximated from life tables of 1933 to 1937.) It may be remarked in passing, however, that the same difficulty arises in converting gross rates based on the total number of children born by duration of marriage to net rates, for the reason that by the time a cohort of married women nears the end of the childbearing period, the duration of their marriage has been from 20 to 25 years.

V. COHORT REPLACEMENT RATES

The calendar year net reproduction rate is criticized on the ground that it projects temporary conditions too far into the future; the generation rate is criticized because it reflects conditions of a number of years past. The objections to both of these rates are partially obviated if a measure is used which reflects the cumulated number of births per 100 women of all childbearing ages up to a particular date, regardless of what age has been attained by that date. That is to say, in 1945 the women born in 1900 had attained age 45; those born in 1901 had attained the age 44; etc. Consequently, only the oldest cohort has completed childbearing; the others have varying periods to complete. If the cumulated reproduction of all reproducing cohorts[8] is set in relationship to a "standard" performance which would result in a net reproduction rate of 100, the result is an approximate measure of the rate of reproduction of all women of all ages up to the date at which the calculation is made.

As the cohort quota replacement rate has not previously been discussed in print, the following notation is introduced in relation to Tables 4, 5, and 6:

P_a = Female generation births per 100 women at age a.

M_a = Female generation survival rate up to age a.

[8] In the calculations presented in Tables 5 and 6, generations 5 years apart are shown instead of those beginning in every calendar year.

M_aP_a = Net reproduction at age a.

n = Number of generations in which an age group reproduces during the period for which the rate is calculated.

$$\overline{M_aP_a} = \frac{\sum M_aP_a}{n} = \text{Mean age-specific net reproduction.}$$

(1) $\sum_{15}^{45} \overline{M_aP_a}$ = Average cohort replacement for all cohorts.

$$M_aP_a{}^0 = \frac{\overline{M_aP_a}}{\sum_{15}^{45} \overline{M_aP_a}} = \text{Age replacement quota.}$$

(2) $\sum_{15}^{a+5} M_aP_a{}^0$ = Cohort replacement quota.

The data for these calculations are arranged exactly as they are in the calculation of the generation net reproduction rate—i.e., by obtaining the female age-specific fertility rate for women of each age in the calendar year when they attained that age (Table 5) and multiply-

TABLE 5

FIVE-YEAR FEMALE BIRTHS PER 100 WOMEN IN EACH AGE—COHORTS BORN 1900–1925* (PA)

Age of Women When Children Were Born	Women Born in: 1900	1905	1910	1915	1920	1925
	Reached Age 15 in: 1915	1920	1925	1930	1935	1940
15–19	13.58	13.58	13.00	11.06	10.91	12.03
20–24	38.63	34.29	30.56	30.75	35.94	
25–29	33.42	29.10	28.28	33.66		
30–34	20.80	19.11	22.07			
35–39	11.45	11.83				
40–44	3.98					

* Based on revised estimates of P. K. Whelpton for under-registration and incompleteness of registration area. U. S. Bureau of the Census, "Forecasts of the Population of the United States," p. 17.

ing by the generation age-specific survival rate from 0 up to that age (Table 4). The results of these calculations are shown in Table 6 under the heading, "Surviving Female Children per 100."

With this arrangement of the data, the next step is to compute the average reproduction in each age ($\overline{M_aP_a}$). These averages are shown

TABLE 6

FIVE-YEAR NATIVE WHITE NET REPRODUCTION STANDARD QUOTAS AND NET REPRODUCTION FREQUENCIES UP TO 1944, FOR COHORTS BEGINNING REPRODUCTION IN 1915–1940

Age of Mother When Child Was Born (a to $a+5$)	Surviving Female Children per 100 (M_aP_a)						Mean Net Reproduction ($\overline{M_aP_a}$)	Age Group Standard Quota ($M_aP_a{}^0$)
	Women Born in:							
	1900	1905	1910	1915	1920	1925		
	Reached Age 15 in:							
	1915	1920	1925	1930	1935	1940		
15–19	11.92	12.22	11.89	10.24	10.11	11.40	11.30	10.79
20–24	34.33	31.20	28.18	28.72	33.89		31.26	29.85
25–29	30.25	26.68	26.29	31.57			28.70	27.41
30–34	18.95	17.61	20.57				19.04	18.18
35–39	10.50	10.94					10.72	10.24
40–44	3.70						3.70	3.53
Observed Cohort Reproduction $\sum_{15}^{a+5} M_aP_a$	109.65	98.65	86.93	70.53	44.00	11.40	104.72	
Cohort Standard Quota $\sum_{15}^{a+5} M_aP_a$	100.00	96.47	86.23	68.05	40.64	10.79	—	100.00
Cohort Replacement $\sum_{15}^{a+5} \frac{M_aP_a{}^0}{M_aP_a{}^0}$	1.096	1.023	1.008	1.036	1.083	1.057	1.047	

in the next to the last column of Table 6. The sum of these averages (Formula 1) yields an average net reproduction rate for all women of all childbearing ages from the time when they were 15 until the date of the calculation.

This calculation may be extended to determine a quota of "normal" reproduction to which each individual cohort may be compared in order to measure its relative performance up to the date of the calculation. These quotas are established as follows: Divide each of the average age-specific reproduction rates ($\overline{M_a P_a}$) by the total of these rates, thus converting the rates for each age into the percentage of births which normally occur in that age. Since the sum of the percentages equals 1, they determine age-specific frequencies which, if equaled, would result in a net reproduction rate of 1; thus, the extent of deviation of actual reproduction from such a quota measures the extent of the deviation of reproduction from a standard stationary rate. These quotas are shown in the last column of Table 6, and are cumulated in the next to the last line of Table 6 for comparison with actual cumulated reproduction. It is thus possible to compare any specific age in any cohort with a corresponding quota, or to compare the cumulated reproduction of each cohort with its quota.

Whereas the major weighting of the experience of a completed generation is 17 years before the end of their experience (when women are age 28), the major weighting of the experience of the total population is about 11 years preceding the date of the calculation. It will be noted that in Table 6, 11 of the 21 cohort-age groups measure fertility for the 10 years immediately preceding the date of the calculation, and 10 of the 21 relate to the older generations whose experience extends from 10 to 30 years back of the date of calculation. A series of such cohort replacement rates is even more stable than the rate for a series of generations.

In avoiding some of the disadvantages of other rates, however, such calculations also lose some of their advantages, since they refer neither wholly to the present nor wholly to the complete experience of generations. However, a series of such calculations extending over a number of years would provide a basis for extrapolation which is more sensitive to present conditions than a series of generation rates.

Whelpton presented fertility data for such incomplete cohorts in a paper as yet unpublished.[9] He did not, however, convert the gross reproduction rates of these incomplete cohorts into net rates.

Paper delivered at the 1947 Session of the International Statistical Institute.

VI. PARITY ADJUSTED RATES[10]

Although Whelpton has published rates adjusted for age, parity, fecundity, and marriage, we shall discuss at this point only the adjustments which were made for parity. His observations led to the conclusion that, if birth rates by order of birth, as calculated for a single abnormal calendar year, are assumed to apply to a cohort of women who actually pass through the childbearing period, impossible results can be obtained. Such calculation for the year 1942 produced the impossible rate of 1,084 first births per 1,000 women. (With the 1947 calendar year as a basis, the results would be still more impossible.)

His first approach to the solution of this difficulty was to standardize for birth order by beginning with the life table cohort of women who had had no children and successively deducting the mothers who bore a first child and survived to the next age from the base of the rate for second births, and so on for each higher order of birth. The resulting parity-adjusted rate had the advantage of avoiding impossible results, but was still open to the objections which have been made to rates which generalize the experience of a single calendar year. Realizing this, Whelpton in some of his later (unpublished) work has made calculations according to order of birth on the basis of the actual recorded generation experience. If recorded statistics are available, this method has all of the advantages and disadvantages of the generation ratio described in the preceding section. The period for which Whelpton has made his later generation birth order calculations is the same as that used for the generation reproduction rates calculated by the author. They are thus susceptible to reduction to net rates by the use of generation life tables and have the obvious advantage of being adapted to studies in which birth order is an important factor singled out for special consideration.

VII. CONCLUSION

The variety of experimentation with rates in all countries where data are available should provide the basis for extending and refining the analysis of fertility. The development of a variety of techniques should, therefore, be welcomed. A review of the foregoing discussion reveals that no one of these rates is most appropriate for all purposes, but that each is well adapted to some particular purpose. Perhaps the effort of some demographers to arrive at a single "optimum" measure which will reflect the condition of the moment and at the same time

[10] Whelpton, P. K., "Reproduction Rates Adjusted for Age, Parity, Fecundity, and Marriage," *Journal of the American Statistical Association*, December 1946.

provide a basis for the measurement of underlying fertility trends is as difficult as it is "to eat one's cake and have it, too." Certainly no measure reviewed in this article is well adapted to both purposes. Demographers who are interested in "invariable fertility trends" might do well to review the history of economic prediction which relies not on one but on the variety of interacting indices.

It will be observed that the rates summarized above fall into two general groups:

(1) Those which use as their basis the experience of a single calendar year and thus concentrate on the condition of the moment. They are subject to violent fluctuations, such as those that have taken place in the last 20 years. These fluctuations make it difficult to select the proper base period from which to project a trend and to determine the shape of the curve which is being projected.

(2) Those which follow the generation technique of recording the actual fertility experience of mothers at a particular age in the year during which they attained that age. These rates include the general generation rate, the rates based on the cumulated number of children born to women who were married at a certain date and age, and the cumulated birth order rates based on actual chronological experience. This latter type is not so sensitive to year-to-year changes, and, consequently, does not reflect conditions of the immediate present as accurately as do rates based on the conventional net reproduction technique. They are, however, more stable and are based on a longer time period, reflecting the impact of past fertility on present rates. Not the least of the advantages of the calculations of the generation type is the fact that they are adapted to conversion to net rates by allowing for generation improvement in mortality.

ON ESTIMATING THE MEAN AND STANDARD DEVIATION OF TRUNCATED NORMAL DISTRIBUTIONS*

A. C. COHEN, JR.
University of Georgia

The problem considered is that of estimating the mean and standard deviation of a normally distributed population from a truncated sample when neither count nor measurements of variates in the omitted portion of the sample is known. Formulas are developed whereby certain special functions required in solutions given for this problem by Karl Pearson and Alice Lee and by R. A. Fisher may be readily evaluated with the aid of an ordinary table of the areas and ordinates of the normal curve. A method of successive approximations is illustrated which, with the aid of the above formulas permits the utilization of either the Pearson-Lee or the equivalent Fisher method to obtain the desired estimates with an improvement in accuracy regardless of whether or not the special tables ordinarily required by these two methods are available.

I. PEARSON-LEE METHOD

KARL PEARSON and Alice Lee [1], and Alice Lee [2] employed the method of moments as early as 1908 to develop formulas which may be used to estimate the mean and standard deviation of a normally distributed population from data provided by a truncated random sample from which all record including both count and measurements of all variates whose value is beyond a given truncation point, has been omitted. Their results except for minor changes in notation may be summarized as follows:

$$m' = x_0' - h'\sigma' \tag{1}$$

$$\sigma' = \frac{1}{n}\sum x \cdot \psi_2 \tag{2}$$

$$\frac{n\sum x^2 - (\sum x)^2}{(\sum x)^2} = \psi_1. \tag{3}$$

In the above equations, m' and σ' are estimates of the mean and standard deviation of the population (complete distribution). x_0' is the point of truncation measured on the original scale of the variate x'. The omitted portion of the sample is here considered to be to the left of x_0'. The summations $\sum x$ and $\sum x^2$ are taken about x_0' as an origin.

* A portion of this paper was presented before the Southeastern Section of the Mathematical Association of America at Tuscaloosa, Alabama, March 19, 1949.

h' is the point of truncation measured in standard units of the population; that is,

$$h' = \frac{x_0' - m'}{\sigma'}.$$

n is the number of variates in the truncated sample. ψ_1 and ψ_2 are moment functions of h'. Tables of both these functions evaluated to three places of decimal at intervals of 0.1 in h' are contained in the original papers and also in "Tables for Statisticians and Biometricians" [3] Vol. I. Table XI and Vol. II. Table XII. To apply the Pearson-Lee method, one proceeds as follows:

(1) Evaluate the left side of Equation (3) from the sample data. Enter the table of ψ_1 with this value and obtain h' by inverse interpolation.

(2) Using the above value of h' as the argument, read ψ_2 from the appropriate table and apply Equation (2) to obtain σ'.

(3) With both σ' and h' determined, then apply Equation (1) to obtain m'.

Unfortunately the tables required for this method are not as widely distributed as might be desired. Furthermore the entries contain too few significant digits and are tabulated at an interval of the argument (0.1) that is too wide to permit h' to be determined with sufficient accuracy for many applications.

II. FISHER METHOD

In 1931, R. A. Fisher [4] demonstrated that the "Maximum Likelihood" estimates for this problem are identical with those obtained by the method of moments. His results, however, were expressed in a slightly different form from those of Pearson and Lee. Fisher employed a moment function of h' (designated by ξ in his discussion) which he labeled as an I_n function and which may be defined as:

$$I_n(h') = \frac{1}{\sqrt{2\pi}} \int_{h'}^{\infty} \frac{(t - h')^n}{n!} e^{-t^2/2} dt \tag{4}$$

and for which the following relations hold:

$$(n + 1)I_{n+1} + h'I_n - I_{n-1} = 0; \qquad n > -1. \tag{5}$$

$$\frac{dI_n}{dh'} = -I_{n-1}. \tag{6}$$

The Fisher results in terms of the I_n functions are:

$$\sigma' = \frac{1}{n} \sum x \frac{I_0}{I_1} \tag{7}$$

and

$$\frac{n \sum x^2}{(\sum x)^2} = \frac{2I_0I_2}{I_1^2} . \tag{8}$$

His equation for obtaining m' after σ' and h' are determined is the same as that given by Pearson and Lee (Equation 1.)

Tables of I_0, I_1 and I_0I_2/I_1^2 (labeled $Hh_0Hh_2/(Hh_1)^2$) as required for use in the Fisher formulas (Equations 7 and 8) are included in "Mathematical Tables" Vol. 1 of the British Association for the Advancement of Science. Entries in these tables are given to a greater number of decimals than the Pearson-Lee tables (from 6 to 9 significant digits for most entries) but the interval of the argument h' (0.1) is the same as for the Pearson-Lee tables. The greater number of significant digits permits greater accuracy in determining h', but only by resorting to inverse interpolation formulas involving the second and higher order differences. At best such computations are rather tedious and somewhat bothersome to carry out. With regard to availability, the B.A.A.S. Tables are perhaps even less widely distributed than the Pearson Tables. The application of the Fisher results is almost identical with that of the Pearson-Lee results. The quantity on the left side of (8) is computed from the sample data and h' is determined by inverse interpolation from the table of I_0I_2/I_1^2. Using the value of h' thus determined, I_0 and I_1 are obtained from the appropriate tables and σ' is computed by use of Equation (7). Equation (1) is then used as before to obtain m'.

III. EQUIVALENCE OF PEARSON-LEE AND FISHER RESULTS

Since Equations (2) and (3) are equivalent to (7) and (8) it follows that:

$$\psi_2 = I_0/I_1 \tag{9}$$

and

$$\psi_1 = \frac{2I_0I_2}{I_1^2} - 1. \tag{10}$$

IV. NEW CONTRIBUTIONS

In the present paper, equations are derived which permit the calculation of ψ_1, ψ_2 and likewise $2I_0I_2/I_1^2$ without resort to any tables other than an ordinary table of areas and ordinates of the normal frequency curve such as can be found in practically any handbook of

mathematical tables. By using the formulas presented herein, the Pearson-Lee or the Fisher technique can be readily applied regardless of the availability of the special tables previously mentioned. Even when the special tables are available, the formulas developed in this paper permit the attainment of greater accuracy in determining h' and consequently in obtaining σ' and m' with a minimum computing effort.

V. DERIVATIONS

Let $n=0$ and 1 respectively in Equation (5) to obtain

$$I_1 = I_{-1} - h'I_0 \tag{11}$$

and

$$2I_2 = I_0 - h'I_1. \tag{12}$$

Let $n=0$ in Equation (6) and we have $dI_0/dh' = -I_{-1}$. If now $n=0$ in Equation (4) it follows that

$$I_0 = \frac{1}{\sqrt{2\pi}} \int_{h'}^{\infty} e^{-t^2/2} dt \tag{13}$$

which is recognized as the area under the normal curve to the right of the ordinate $t=h'$. Direct differentiation of (13) gives $dI_0/dh' = -\phi(h')$ where $\phi(h')$ is the ordinate of the normal curve at $t=h'$.

$$\left(\phi(h') = \frac{1}{\sqrt{2\pi}} e^{-h'^2/2}\right).$$

Consequently we may write

$$I_{-1}(h') = \phi(h'). \tag{14}$$

Upon substituting the results of Equations (11), (12) and (14) in the right side of Equation (8) it follows that

$$\frac{2I_0I_2}{I_1^2} = \frac{[I_0 - h'(\phi - h'I_0)]I_0}{[\phi - h'I_0]^2}. \tag{15}$$

Now if we define

$$Z(h') = \frac{\phi(h')}{I_0(h')}, \tag{16}$$

and divide both the numerator and denominator of the right side of (15) by I_0^2 we obtain

$$\frac{2I_0I_2}{I_1^2} = \left(\frac{1}{Z - h'}\right)\left(\frac{1}{Z - h'} - h'\right). \tag{17}$$

From Equation (10) it then follows that

$$\psi_1 = \left(\frac{1}{Z - h'}\right)\left(\frac{1}{Z - h'} - Z\right). \tag{18}$$

Similarly from Equations (11) and (14) we obtain

$$\frac{I_0}{I_1} = \psi_2 = \frac{1}{Z - h'}. \tag{19}$$

Equations (17), (18) and (19) are thus expressed in convenient forms to permit rapid calculation of the special functions required by both the Pearson-Lee and the Fisher methods for any values of the argument h' as may be necessary, provided only that existing tables of the areas and ordinates of the normal frequency curve are available. These formulas would also be useful in extending both the Pearson-Lee and the Fisher tables by computing additional entries at closer intervals of the argument h'.

VI. DETERMINING h' BY SUCCESSIVE APPROXIMATIONS

Although tables are used in both the Pearson-Lee and the Fisher methods, the basic problem involved is the solution of either Equation (3) or its equivalent Equation (8) for h'. A method of successive approximation which makes use of (17), (18), and (19) and simple linear interpolation has been found to be quite satisfactory for determining h' to additional significant digits. If either the Pearson-Lee or the Fisher tables are available they might be used to furnish a first approximation of h'. For use when neither of these tables are available, a graph of ψ_1 for values of h' and from -3.5 to $+3.5$ is given in Figure 1. In the absence of both tables and graph a reasonably satisfactory first approximation to h' can be obtained from

$$h' \sim (x_0' - \bar{x})/s_x \tag{20}$$

where $\bar{x}$ is the mean and s_x is the standard deviation of the truncated sample. To compute s_x use the formula

$$(ns_x)^2 = n\sum x^2 - (\sum x)^2.$$

VII. NUMERICAL EXAMPLE

The various steps involved in determining h' and subsequently σ' and m' by the successive approximation technique mentioned in the previous paragraph can best be understood by computing these quantities for a typical set of data such as the following:

$$n = 37;\ x_0' = 0.850;\ \sum x = 51.8600;\ \text{and}\ \sum x^2 = 98.0156$$

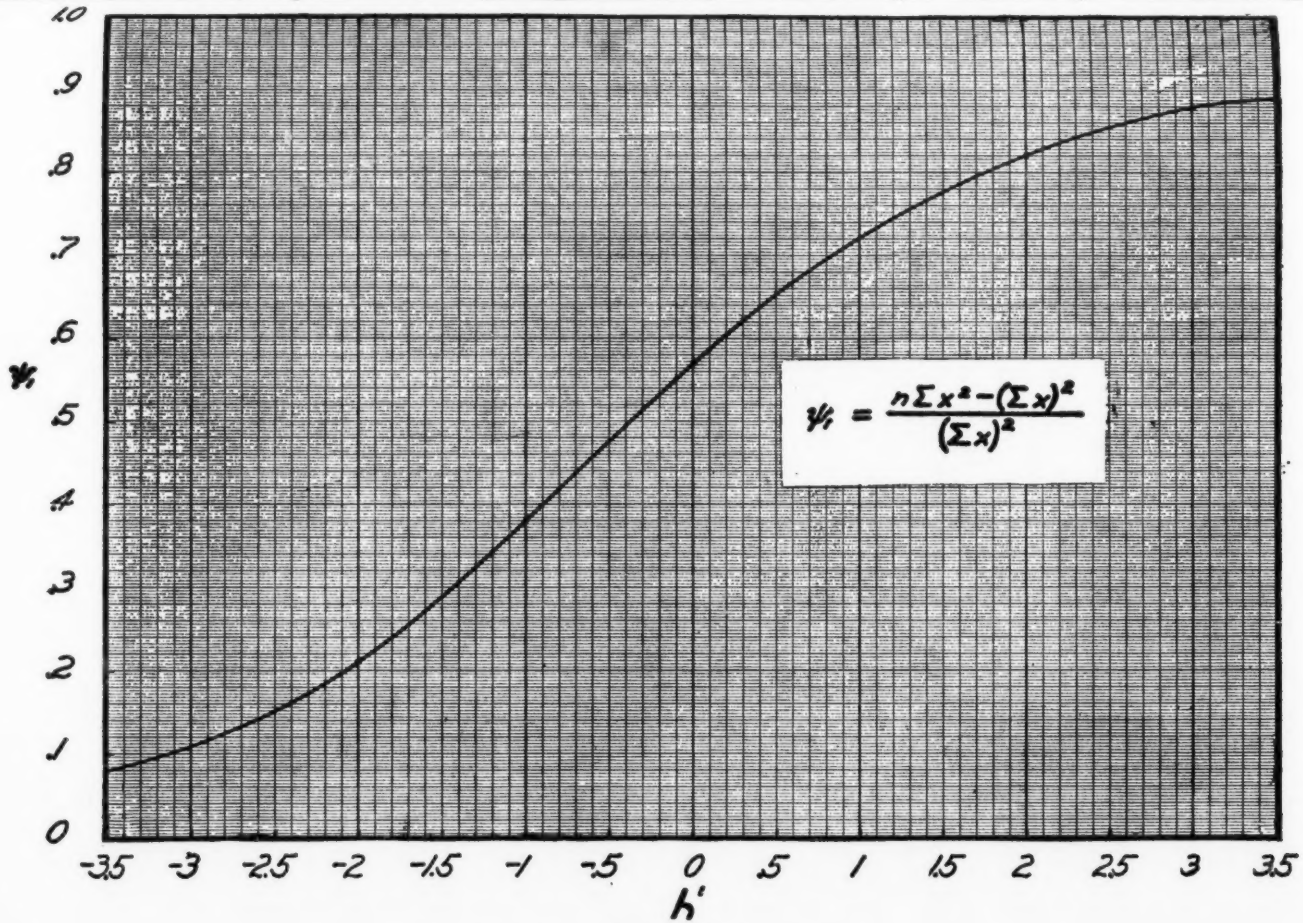


FIGURE 1
GRAPH SHOWING RELATION BETWEEN ψ_1 AND h'

where x has been measured from the terminus as an origin ($x_i = x_i' - x_0'$). Thus $\bar{x} = 1.401622$; $s_x = 0.83$; and

$$\frac{n\sum x^2 - (\sum x)^2}{(\sum x)^2} = 0.348441.$$

From the graph of Figure 1, we find $h_1' = -1.20$ as a first approximation. By direct substitution in Equation (18) it is found that $\psi_1(-1.20) = 0.341734$. Since this value is less than 0.348441 and since ψ_1 is an increasing function of h', it is necessary to select a value greater than -1.20 as a next approximation. We then find that $-1.20 < h' < -1.10$. Linear interpolation gives $h_2' = -1.16$ as a closer approximation and following the next step we establish that $-1.165 < h' < -1.164$. Again linear interpolation gives $h_3' = 1.1643$ which is accepted as a final approximation since this determination is sufficiently accurate for our purposes.

The tables of areas and ordinates of the normal distribution used in making these computations contained six significant digits and were tabulated at intervals of 0.01 for the argument. In using these tables it was necessary to employ interpolation for obtaining entries only for the two final approximations. Table 1 details the computations involved in the various steps described above.

TABLE 1

h'	I_z	ϕ	Z (3) ÷ (2)	$Z - h'$ (4) − (1)	ψ_z 1/(5)	$\frac{1}{Z-h'} - Z$ (6) − (4)	ψ_1 (6) × (7)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
−1.10	.864334	.217852	.252046	1.352046	.739620	.487574	.360619
−1.20	.884930	.194186	.219436	1.419436	.704505	.485069	.341734
−1.16	.876976	.203571	.232128	1.392128	.718325	.486197	.349247
−1.17	.879000	.201214	.228912	1.398912	.714841	.485929	.347362
−1.164	.877786	.202628	.230840	1.394840	.716928	.486088	.348490
−1.165	.877988	.202392	.230518	1.395518	.716580	.486062	.348302

It has been the writer's experience that by systematically arranging computations as shown in the above table the tedium of making each individual determination of $\psi_1(h')$ is considerably reduced.

By interpolation from Table 1 it is readily found that

$$(Z - h')\big|_{h' = -1.1643} = 1.395043$$

and it follows from Equations (2) and (19) that

$$\sigma' = 1.401622/1.395043 = 1.0047.$$

From Equation (1) it then follows that $m' = 0.850 - (1.0047)(-1.1643) = 2.020$. This completes the solution.

VIII. VARIANCE OF ESTIMATES

R. A. Fisher obtained the following formulas for the variance of σ and h' (Fisher's ξ)

$$V(\sigma) = \frac{\sigma^4 \mu_2}{n\{\mu_2'\mu_2 + \sigma^2(2\mu_2 - \mu_2')\}}, \tag{21}$$

$$V(h') = \frac{\sigma^2(\mu_2' + \sigma^2)}{n\{\mu_2'\mu_2 + \sigma^2(2\mu_2 - \mu_2')\}}, \tag{22}$$

and for the correlation between the sampling errors of σ and h'

$$r_{\sigma,h'} = \frac{+\sigma\mu_1'}{\sqrt{\mu_2(\mu_2' + \sigma^2)}}, \tag{23}$$

where μ_k' is the kth moment of the truncated sample about its terminus x_0' and μ_k is the kth central moment of the truncated sample.

It can be shown that in terms of the function Z defined by Equation (16) that the above equations become

$$V(\sigma) = \frac{\sigma^2}{n}\left[\frac{1 - Z(Z - h')}{[1 - Z(Z - h')][2 - h'(Z - h')] - (Z - h')^2}\right], \tag{24}$$

$$V(h') = \frac{1}{n}\left[\frac{2 - h'(Z - h')}{[1 - Z(Z - h')][2 - h'(Z - h')] - (Z - h')^2}\right], \tag{25}$$

$$r_{\sigma,h'} = \frac{+(Z - h')}{\sqrt{[1 - Z(Z - h')][2 - h'(Z - h')]}}. \tag{26}$$

REFERENCES

[1] Pearson, Karl and Lee, Alice, "On the Generalized Probable Error in Multiple Normal Correlation," *Biometrika*, Vol. 6, 1908, pp. 59–68.

[2] Lee, Alice, "Table of Gaussian 'Tail' Functions When the 'Tail' is Larger Than the Body," *Biometrika*, Vol. 10, 1915, pp. 208–215.

[3] Pearson, Karl, editor (1914, 1931). *Tables for Statisticians and Biometricians*, Part I (1914, 3rd edition 1930) and Part II (1931), Cambridge University Press.

[4] Fisher, R. A., "Properties and Applications of *Hh* Functions," Introduction to *Mathematical Tables*, Vol. 1, British Assn. for the Advancement of Science, 1931, pp. xxvi–xxxv.

ON SOME MATHEMATICAL PROBLEMS ARISING IN THE DEVELOPMENT OF MENDELIAN GENETICS*

HILDA GEIRINGER
Wheaton College, Massachusetts

In this paper some basic problems of theoretical Mendelian genetics are presented. The general situation of an arbitrary number of linked loci for "diploids" (normal organisms) as well as for "autopolyploids" is considered, under random mating. Particular stress is laid upon formulating the problems and results within the framework of probability theory, expressing them in terms of three basic distributions, the distributions of genotypes and of gametes, and the segregation distribution. The task of the mathematical theory consists in establishing recurrence relations for the distributions of gametes and of genotypes, in integrating them and in investigating the limit behaviour of those distribution as n, the number of discrete non overlapping generations tends toward infinity.

IN A PAPER delivered in 1935 at the Massachusetts Institute of Technology, J. B. S. Haldane [9, g] spoke on "Some Problems of Mathematical Biology." In this lecture, which covers a much broader subject than the present paper, he suggests the following classification of the problems of mathematical biology. A first group of problems is concerned with the life of the cell; next, one may consider the analysis of a whole organism composed of cells, like a tree or an animal; on a third level, one studies the mutual relations of a number of organisms, —members of the same brood, individuals of the same population,— and investigates the biological fate of such populations.

The problems dealt with in the present paper belong to the third group, which is concerned with entire populations. The mathematical tools for these investigations are offered by probability theory and statistics, since this is the branch of mathematics which deals with mass phenomena. However, we will not consider here the manifold problems connected with the statistical evaluation of biological observations. From a systematic standpoint problems of biological statistics are not different from the statistical problems occurring in connection with other series of observations. The more specific task of mathematical genetics is comparable to what is done in the kinetic theory of gases, in the theory of quanta, etc. One starts by formulating probability laws that express certain biological facts and then tries to deduce

* Lecture delivered January 1948 at the University of Chicago.

in a mathematical way consequences which are to be checked by observations. The simplest example of such a procedure presents itself in Mendel's theory of the heredity of a single character.

1. RANDOM INHERITANCE OF ONE CHARACTER. THE BASIC PROBABILITY DISTRIBUTIONS

Mendel [17] studied certain single traits or characters, like the heredity of the color of the flower of peas or the size (tall or short) of this plant, by actually counting the numbers of each type of progeny which resulted from a given cross, thus applying statistics to the phenomena of inheritance. His fundamental discoveries were made on the garden pea, *Pisum sativum*. Systematic observations and analysis led him to the following daring hypothesis. The visible color (red or white in case of the peas,—red, white or pink in case of four-o'clocks) is dependent on a pair of factors which we may denote for the moment by R and W. Each fertilized zygote (and each cell of the organism into which it develops) contains *two* of these factors and may thus be of the type RR, or RW, or WW. The gametes (egg and sperm) contain one factor only, selected at random from among the two factors contained in the cell out of which the gamete is formed. A new zygote of the following (filial) generation is then formed by the random union of two gametes, and, consequently, again contains two factors, and so on.

Such factors have been called Mendelian factors, unit factors, or *genes*. There is no doubt, today, that the genes are located in the *chromosomes*. Two genes alternative to one another, like R and W, are spoken of as allelomorphic factors, or *alleles*. There may be more than two members of such an "allelomorphic series" e.g. t=tall, s=short, and d=dwarf with respect to size. A zygote which contains two alleles of the same kind, like RR, or WW, or tt, etc., is said to be *homozygous* for the factor in question while an RW-plant or a plant of type td, is called *heterozygous* or *hybrid*. *Locus* means the particular place in a particular chromosome at which there are alleles. *Character* relates to the effect of genes. A single character, like color, may be affected by many loci and, on the other hand, the same locus may influence several visible characters. In the case of the flower-color of peas, however, this character is determined by *one* locus.

In the first paragraph of this section we have presented the essential content of Mendel's famous *first law*. Let us attempt a more mathematical formulation. We assume that, corresponding to each locus, there exists a *random variable*, x, which may take on r distinct values,

the r alleles, $x = a_1, a_2, \cdots, a_r$. In the example of the color of peas, $r = 2$; here a_1 stands for "red" and a_2 for "white." *Three* alleles determine the human blood groups, and there are at least $r = 14$ known alleles for the eye color of *Drosophila melanogaster*, the small vinegar fly which plays as important a role in modern genetics as did Mendel's peas a century ago. With respect to such a locus the genetic type of an organism (zygote as well as grown organism) is specified, not by one, but by *two* values, x and y, of this random variable which may or may not be equal to each other. They represent the organism's maternal and paternal heritage, since it receives one allele from its mother and the other from its father.

We assume that two organisms are genotypically the same with respect to the locus in question if in one of them x comes from the mother and y from the father, while in the other organism it is the other way round. If we denote the type of an organism by (xy), where the first letter denotes the maternal and the second the paternal heritage, this assumption reads,

$$(1) \qquad (xy) = (yx).$$

Consequently there are $r(r+1)/2$ possible genotypes in this case. If $r = 2$ the three types are often denoted by (AA), (Aa), and (aa). Next we suppose *distinct, non-overlapping generations.* A basic assumption is then that in a certain "initial" generation the possible types (xy), $(x = a_1, \cdots, a_r)$; $(y = a_1, \cdots, a_r)$ are distributed according to a *law of probability.* This distribution will be called the initial probability *distribution of genotypes,* $w^{(0)}(xy)$. From it, by means of the hypotheses which characterize our problem, the distributions of genotypes for later generations will follow. We denote the distribution for the nth generation by $w^{(n)}(xy)$. In accordance with (1) we must assume that

$$(2) \qquad w^{(n)}(xy) = w^{(n)}(yx) \qquad (n = 0, 1, \cdots), \left(\begin{matrix} x \\ y \end{matrix} = a_1, \cdots, a_r\right).$$

There is no loss of generality if we suppose that the initial distribution of genotypes is the same for males and females. In fact, it is easily seen that under random mating any difference between initial distributions disappears in the first filial generation. For this distribution we have

$$(3) \qquad \sum_x \sum_y w^{(n)}(xy) = 1, \qquad (n = 0, 1, \cdots), \left(\begin{matrix} x \\ y \end{matrix} = a_1, \cdots, a_r\right).$$

Next Mendel assumed that in the formation of a new individual the

parent (through the gamete) transmits to the offspring one of its two genes, either x or y. The choice between the two values happens according to a probability law. Mendel's assumption, deduced from and confirmed by observation, is that *the two probabilities* for the segregation of either of the two genes *are equal*. With a view to more general cases, we introduce a second basic distribution, the *segregation distribution*. Denote by l_0 the probability that the paternal gene be transmitted, by l_1 the analogous probability for the maternal gene. Then, in the particular case under consideration

$$l_0 + l_1 = 1, \qquad l_0 = l_1 = \tfrac{1}{2}. \tag{4}$$

The segregation distribution is not so trivial in more general situations. We shall however retain two of the assumptions made here: (a) The segregation distribution is independent of "n", it does not change through the generations. (b) It does not depend on the genotype of the parent. We shall see that the segregation distribution plays a basic role wherever random mating is considered. Certain problems where selection is involved cannot be described in terms of the segregation distribution. (See section 2.)

From these two distributions we derive the third important distribution, the *distribution of gametes*, $p^{(n)}(x)$, $(n=0, 1, \cdots)$, $(x=a_1, \cdots, a_r)$. This is the probability that the gamete be of constitution x, i.e. possesses the gene x. Let us compute this distribution. In order to transmit the gene x the parent must possess this gene and transmit it. Accordingly we have for example for the gene $x=1$, (writing $1, 2, \cdots, r$ instead of $a_1, a_2, \cdots, a_r$), $p^{(n)}(1)=w^{(n)}(11)(l_0+l_1)+w^{(n)}(12)l_1+w^{(n)}(21)l_0+\cdots+w^{(n)}(1r)l_1+w^{(n)}(r1)l_0$. This formula is easily understood. For example, the term $w^{(n)}(12)l_1$ is the probability that the parent possesses the maternal gene "1" and the paternal gene "2," multiplied by the probability, l_1, that the transmitted gene be the maternal gene. Obviously the sum of all such terms, as contained on the right side of the preceding formula, gives the probability of a gamete of type "1," (if we assume certain circumstances which we shall analyze presently). Because of (2) and (4) the formula may be written as follows

$$p^{(n)}(x) = \sum_{y}^{1\cdots r} w^{(n)}(xy) \qquad (x = 1, \cdots, r, \ n = 0, 1, \cdots) \tag{5}$$

with

$$\sum_{x}^{1\cdots r} p^{(n)}(x) = 1. \tag{5'}$$

A new individual of the $(n+1)$st generation is formed by the fusion of two gametes of the nth generation. The assumption that this fusion happens at random is expressed in the formula:

$$w^{(n+1)}(xy) = p^{(n)}(x)p^{(n)}(y) \qquad \begin{pmatrix} x \\ y \end{pmatrix} = 1, 2, \cdots, r \tag{6}$$

and that finishes the cycle, since we now know $w^{(n+1)}(xy)$, the distribution of genotypes in the next generation.

Before analyzing the assumptions which underlie these formulas I want to derive the famous *law of constancy of gametic proportions*, first recognized by the biologist W. Weinberg [25] and proved by the mathematician G. H. Hardy [10]. We obtain from (5) and (6):

$$\begin{aligned} p^{(n+1)}(x) &= \sum_{y}^{1\cdots r} w^{(n+1)}(xy) \\ &= \sum_{y}^{1\cdots r} p^{(n)}(x)p^{(n)}(y) = p^{(n)}(x) \sum_{y}^{1\cdots r} p^{(n)}(y) \\ &= p^{(n)}(x) \qquad (n = 0, 1, \cdots). \end{aligned} \tag{7}$$

From this it follows that

$$w^{(n+1)}(xy) = w^{(n)}(xy), \qquad (n = 1, 2, \cdots). \tag{7'}$$

These basic results state that *the distribution of gametes remains constant throughout, while the distribution of genotypes remains constant from the first filial generation on.* These results hold in this form in our "simplest case" only. In more complicated problems the first result must, in general, be modified and the second fails.

2. REMARKS ON "RANDOM MATING"

Let us try to give a definition, or rather a mathematical explanation, of the concept of random mating as opposed to "selection." I do not propose to discuss the meaning in probability of the concepts "random" or "randomness" or "random variable," which are inherent in any probability theory. Although there is much controversy with respect to these notions, every statistician attributes a certain meaning to them.

It seems to me that the mathematical meaning of "random mating" is contained in the equations (5) and (6) by which $w^{(n+1)}$ is derived from $w^{(n)}$. In an equation of type (5), the distribution of gametes in the nth generation appears as a linear expression in the $w^{(n)}(xy)$, the values of the distribution of zygotes, with constant coefficients which

are sums of values of the segregation distribution. An equation of type (6) expresses the random fusion of the two gametes. In the various cases of selection and mutation *one or more of the assumptions which lead to* (5) *and* (6) *cannot be maintained.* We shall see, for example, that these equations contain the assumption that within the population there are no genotypic differences with respect to the attainment of maturity, to fecundity, or to mortality (no matter whether such differences are "natural" or "planned"). Another assumption contained in our equations is that the choice of the mate happens at random. Two examples will illustrate these points.

Consider, as in section 1, the "simplest case" of one locus and assume $r=2$ alleles. Also suppose that the choice of the mate is still due to chance, but that there is a differential viability for the three genotypes. How do we express this mathematically?

Write A and a for the two alleles and introduce for the sake of brevity

$$w^{(n)}(AA) = p_n, \qquad w^{(n)}(Aa) = q_n, \qquad w^{(n)}(aa) = r_n.$$

These p_n, q_n, r_n constitute the distribution of zygotes as generated by their parents. In case of a different viability of the types we can no longer state that the distribution of genotypes at the time of maturity is the same as at time of birth. We have to introduce a *different* distribution, p_n', q_n', r_n', the distribution of the parents of the next generation. Various types of selection as indicated above may be accounted for in this way. We define p_n', q_n', r_n' by means of "selection coefficients" α, β, γ (only the proportions of these three numbers matter) $p_n':2q_n':r_n' = \alpha p_n : 2\beta q_n : \gamma r_n$, with $p_n'+2q_n'+r_n'=1$. The distribution of gametes is now deduced from the p_n', q_n', r_n', rather than from the p_n, q_n, r_n. Writing x_n and y_n instead of $p^{(n)}(A)$ and $p^{(n)}(a)$ the formula which takes the place of (5) is:

$$x_n = p_n' + q_n', \qquad y_n = q_n' + r_n'.$$

Although (5) no longer holds, the rule expressed in (6) is still valid:

$$p_{n+1} = x_n^2, \qquad q_{n+1} = x_n y_n, \text{ etc.}$$

Mathematically this is a problem of a different character, with essentially different results from those of the problem of section 1.

We quote a second example where the random choice of the mate no longer applies and differential viability or an equivalent condition is not assumed. The strongest deviation from the concept of random choice of the partner presents itself if the genotype of the mate is uniquely determined by the individual's own genotype. Assume that

there exists the rigorous "rule" (natural or artificial) that only identical types may mate. Then there are no longer six but only three possible types of matings: $AA \times AA$, $Aa \times Aa$, and $aa \times aa$. With our original notation we have:

$$w^{(n+1)}(AA) = w^{(n)}(AA) + \tfrac{1}{2}w^{(n)}(Aa),\ w^{(n+1)}(Aa) = \tfrac{1}{2}w^{(n)}(Aa),\ \cdots.$$

The third equation is similar to the first, with A and a interchanged. These equations replace (5) *and* (6). To understand the new situation we write for the right side of the first of these equations: $w^{(n)}(AA)\cdot 1\cdot 1 + 2w^{(n)}(Aa)\cdot 1\cdot \tfrac{1}{4}$. Consider e.g. the second of these two terms. Here $2w^{(n)}(Aa) = w^{(n)}(Aa) + w^{(n)}(aA)$ is the probability of a female of type (Aa). To get the probability of the mating $Aa \times Aa$ we no longer multiply $2w^{(n)}(Aa)$ by $2w^{(n)}(Aa)$, as we would in random mating, but by "1," which represents the *conditional probability* that under the considered "law" the male partner be of type Aa if the female is of type Aa. Hence $2w^{(n)}(Aa)$ is the probability of the mating $Aa \times Aa$. Finally $\frac{1}{4}$ stands for the probability that the offspring of the mating $Aa \times Aa$ be of type AA. More generally, we may denote by $\pi_{\kappa\lambda}$ the probability of a union of the types κ and λ and by $\rho_{\kappa\lambda}^{\mu}$ the probability that the offspring of a mating (κ, λ) be of type μ. We assume that for all κ, λ: $\rho_{\kappa\lambda}^{\mu} = \rho_{\lambda\kappa}^{\mu}$ and $\sum_{\mu}\rho_{\kappa\lambda}^{\mu} = 1$. Here $\rho_{\kappa\lambda}^{\mu}$ need not be, as in our example, the product of two segregation probabilities. It may depend on the types of both parents, whereas our segregation distribution refers to each parent separately and does not depend on the parental types (see sec. 4).

We do not attempt to present a scheme which would apply to the most general case of "selection." The preceding examples are intended to show how new situations call for new concepts.

3. SOME MATHEMATICAL PROBLEMS IN MENDELIAN GENETICS

We now consider some important generalizations of our "simplest problem." In this simplest case there exists no "recurrence problem," since after *one* good mixing the characteristic distributions do not change any more. This is not a typical result. However, the *concepts* introduced for this simplest case remain valid except for certain generalizations.

We continue to assume random mating. As an important example of a more general situation we first consider *various characters* of an individual simultaneously, as e.g. color of the flower, length of the stem, and shape of the seed, etc. The genetic constitution of the individual is still determined by the two gametes which represent its maternal

and paternal heritage. Each gamete, however, no longer consists of a single gene but of a *set of genes*. To denote a genotype let us separate by a semicolon the symbols which stand for its maternal and paternal heritage. Thus a genotype may be denoted by $(x; y)$ where the letter before (after) the semicolon denotes the individual's maternal (paternal) heritage. Also introduce $w^{(n)}(x; y)$, the distribution of genotypes, and assume that, as in (1) and (2):

$$(1') \qquad (x; y) = (y; x)$$

and

$$(2') \qquad w^{(n)}(x; y) = w^{(n)}(y; x).$$

This differentiation with respect to the maternal and paternal heritage is of great importance for the understanding of linkage. It relates however merely *to the formation of new gametes.*

In the example under consideration where we are concerned with the study of m loci, "x" stands for the "maternal" genes, $x_1, \cdots x_m$, while "y" symbolizes the m "paternal" genes, $y_1, \cdots y_m$. The kinds of gametes which this organism may produce depend on the possible combinations of the material it has inherited. These combinations happen according to a probability law which we again call the segregation distribution. It is one of the main tasks of the theory to define a segregation distribution which corresponds to given biological conditions. In a recent paper [5, d], R. A. Fisher says: "The laws of inheritance are the rules whereby, given the constitution of an organism, the kinds of gametes it can produce and their relative frequencies can be predicted." Next, the distribution of gametes $p^{(n)}(z)$ is derived from the distribution of genotypes by means of the segregation distribution. Finally, the random fusion of the two gametes, expressed in the formula $w^{(n+1)}(x; y) = p^{(n)}(x)p^{(n)}(y)$, completes the cycle of inheritance.

It is easy to indicate a few mathematical problems which present themselves in this and in similar situations:

a) The possible genotypes are to be completely and simply enumerated under consideration of the specific biological conditions which characterize a given situation. The same problem exists for the various kinds of gametes.

b) An adequate segregation distribution is to be defined.

In a), as well as in b), the mathematical approach can help to simplify and clarify the concepts. As in other sciences the actual situation in nature is often so complex that "models" have to be constructed representing a compromise between mathematical simplicity and biologi-

cal adequacy. By such a procedure we introduce for example in mechanics models like the "rigid body," the "elastic body," the "ideal fluid," etc.

c) The distribution of gametes, $p^{(n)}(z)$ is to be derived from $w^{(n)}(x; y)$ by means of the segregation distribution, and *recurrence relations* for the $p^{(n)}(z)$ are desired. They are in general, simpler than those for the $w^{(n)}(x; y)$; under random mating both are equivalent.

d) We wish to *integrate* these recurrence equations which are in general, non linear difference equations with constant coefficients; that means we want to express $p^{(n)}(z)$ in terms of "n," of the initial values $p^{(0)}(z)$,—which in turn have to be derived from the $w^{(0)}(x; y)$,—and of the parameters involved in the definition of the segregation distribution.

e) We want to know whether one or more *states of equilibrium* prevail for $p^{(n)}(z)$ and for $w^{(n)}(x; y)$ and if so under what conditions. To settle this we also have to investigate the *limit behaviour* of $p^{(n)}$ and $w^{(n)}$ as "n" tends towards infinity.

In the following we shall comment on the problems a)–e) in the case of: 1) *m loci* (*linkage*), assuming autosomal genes and "diploids"; 2) so-called *autopolyploids* (as opposed to "diploids") for one as well as for several loci.

4. LINKAGE IN THE CASE OF m LOCI AND r ALLELES

As in the first example of the preceding section assume m loci. Mendel's original assumption was that all possible gametic combinations are equally probable. This assumption of *independent assortment* does not introduce any parameters since each of the 2^m possible combinations appears with probability $1/2^m$.[1]

This simple conception was shaken by the observation of "linkage" associated with the names of Bateson and Punnet, of Morgan, Sturtevant, and of many other well known biologists of our time. It appeared that not all possible gametes occur with equal frequency. Consider for a moment the case $m=2$, and r alleles. We denote a type by $(x_1x_2; y_1y_2)$ where x_i as well as y_i, $(i=1, 2)$, relate to the ith locus and each of them takes on the values 1, 2, $\cdots$ r. The four possible types of gametes are then x_1x_2, x_1y_2, y_1x_2, y_1y_2, since one gene is transmitted with respect to each locus. In Mendel's conception each of these four possibilities has the probability $\frac{1}{4}$. In modern genetics these four probabilities are

[1] It is referred to by many authors as "random segregation." This term should not be confused with the general use of the term "random" underlying each probability distribution and not merely the particular one where all probabilities are equal. Such a probability distribution is often called "uniform," in probability calculus.

assumed to be $(1-c)/2$, $c/2$, $c/2$, $(1-c)/2$, where c is the so-called "recombination value." It is often assumed by biologists that c is less than $\frac{1}{2}$. This implies that the genes that came in together exhibit a tendency to stay together; mathematically, however, c may have any value between zero and one. The introduction of the parameter c corresponds to assuming a certain influence of the grand-parents. The case, $m=r=2$, has been completely investigated by Weinberg [25], Robbins [20] and Jennings [13, b, c].

In the general case there are $m(m-1)/2$ recombination values, c_{ij}. These may be defined in a way which is independent of additional assumptions like "chiasma theory," or "linear theory": c_{ij} $(i=1, \cdots m, j=1, \cdots m, i \neq j)$ is the probability that those transmitted genes whose subscripts are i and j come from different parents, or, in other words, the probability that either x_i and y_j or x_j and y_i be transmitted, no matter what happens to the $(m-2)$ other factors. It follows that c_{ij} is a marginal probability of our segregation distribution. It will be seen that, unless we admit some additional biological hypothesis, the $m(m-1)/2$ recombination values are not sufficient for the description of general linkage.

In case of m loci a genotype is characterized by *two m-dimensional vectors*, x_i and y_i $(i=1, \cdots m, x_i=1, \cdots r, y_i=1, \cdots r)$. This simple idea, absolutely basic in the author's approach, is not generally accepted. Often (the consideration being limited to two or three alleles) a type is simply denoted by m pairs of numbers, which is all right only if independent assortment is assumed. Then there are just nine types for $m=r=2$ and in general $[r(r+1)/2]^m$ types. For the understanding of linkage, however, the assumption of two sets, each of m numbers, is necessary. This amounts to distinguishing for example between $(AB; ab)$ and $(Ab; aB)$; in fact, these two types *are* different with respect to the segregation of gametes if $c \neq \frac{1}{2}$, since the first transmits AB with probability $(1-c)/2$, and the second transmits AB with probability $c/2$. There are in this case ten and not nine types and in the general case there are $\frac{1}{2}r^m(r^m+1)$ types.

Thus, a type is denoted by $(x_1, \cdots x_m; y_1, \cdots y_m)$. In the formation of a gamete a new set of m elements is composed in such a way that corresponding to each of the m subscripts either the x-value or the y-value is chosen. Consequently there are 2^m possibilities. The gamete will consist of some of the x_i and of some of the y_i. By fusion of this gamete with another, a zygote is formed which contains in a new combination some of the material found in the parents. This new combination, again, will not persist: when the new being forms sex cells the

new gamete will exhibit a new characteristic combination and the old combinations disappear. Thus the genes recur from generation to generation passing through many individuals which they determine (with respect to certain properties). The combinations change, but not the constituent genes. (This is a very simplified scheme which does not even cover mutations.)

To describe adequately the 2^m possible assortments we introduce a segregation distribution which, in this case, may be called a *linkage distribution*. Denote by S the set of numbers, $1, 2, \cdots m$, by T any subset of S, $(O \leqq T \leqq S)$ and by l_T *the probability that the maternal genes belonging to* T *and the paternal genes belonging to* $T' = S - T$ *be transmitted.* Note that this definition is independent of the genotype of the parent. There are 2^m such probabilities. For obvious symmetry reasons we must assume

$$l_T = l_{T'}. \tag{7}$$

Hence, since the sum of all probabilities is *one*, we have introduced

$$M = 2^{m-1} - 1$$

parameters. For $m \geqq 4$, $M > m(m-1)/2$; so there are in general more linkage-parameters than recombination values. It is desirable to describe linkage in terms of the $m(m-1)/2$ recombination values, that is, to express the M linkage parameters by these values. This problem is closely connected with the so-called "linear arrangement" of the genes of the same linkage group and with the concept of "distance" of linked genes. R. A. Fisher [5, e] recently offered a very suggestive solution. An older solution, fairly generally accepted (but in the author's opinion open to criticism) is due to Haldane [9, c]. Whatever the relation between linkage distribution and recombination values, the linkage distribution must be assumed known if we wish to study the heredity problem of linked genes.

5. SOLUTION OF THE LINKAGE PROBLEM

For the above linkage problem, the author has solved the problems indicated in section 3. We have considered a) the enumeration of the possible genotypes under linkage and b) the definition of the linkage distribution. Before continuing, some important particular forms of the linkage distribution should be specified: 1) In the case of *independent assortment* (random segregation), the 2^m values of the linkage distribution are all equal to each other, hence each equal to $1/2^m$. 2) If there are

S distinct "*linkage groups*," the m-dimensional linkage distribution[2] resolves into the product of S probability distributions. 3) The m loci may form several groups of *completely linked loci*. Then, all recombination values within each group are zero while all recombination values between members of any two different groups equal $\frac{1}{2}$.

The main problem is the *recurrence problem* (problem c) of sec. 3) which always occupies the central place. In the case of m arbitrarily linked (autosomal) genes it has been solved [7a], [13b]. To explain the solution we need the concept of a "marginal distribution," well known to statisticians. If $p(z_1, z_2, \cdots z_m)$ denotes a discrete (arithmetical) probability distribution in m variates, then $p_{12}(z_1z_2) = \sum_{z_3} \cdots \sum_{z_m} p(z_1z_2z_3 \cdots z_m)$ is the probability of the result (z_1z_2) and similarly $p_{ij}(z_iz_j)$, $p_i(z_i)$ or $p_{ijk}(z_iz_jz_k)$ may be defined. Note that there are marginal distributions of "order 1," of "order 2," etc. Obviously we may write $p_T(z_T)$ for $p_{ijk}(z_iz_jz_k)$ if T denotes the set (i, j, k) and z_T the set $(z_iz_jz_k)$. If, finally, we write $p^{(n)}(z)$ for $p^{(n)}(z_1z_2 \cdots z_m)$ our recurrence formula reads

$$p^{(n+1)}(z) = \sum_{(T)} l_T p_{T'}(z_T) p_T;(z_{T'}) \tag{9}$$

where the sum is over all subsets T of S and $p_T(z_T)$ is the marginal distribution whose subscripts are the points of T. We have e.g. (anticipating the result (11):

$$m = 2: \quad p^{(n+1)}(z_1z_2) = 2l(00)p^{(n)}(z_1z_2) + 2l(01)p^{(0)}(z_1)p^{(0)}(z_2)$$

$$\begin{aligned} m = 4: \; p^{(n+1)}(z_1z_2z_3z_4) = 2\{ & l(0000)p^{(n)}(z_1z_2z_3z_4) \\ & + [l(1000)p^{(0)}(z_1)p^{(n)}(z_2z_3z_4) \\ & + \cdots [l(0001)p^{(0)}(z_4)p^{(n)}(z_1z_2z_3)] \\ & + [l(1100)p^{(n)}(z_1z_2)p^{(n)}(z_3z_4) \\ & + l(1010)p^{(n)}(z_1z_3)p^{(n)}(z_2z_4) + \cdot]\}. \end{aligned}$$

In our recurrence formulas the values of the linkage distribution act as essential "separators" between meaningful groups of probabilities. This clear and simple recurrence relation is characteristic for "random mating" and "chromosome segregation" (see also next section).

The next problem consists in the *solution of the recurrence equations.*

[2] A definition of the linkage distribution, completely equivalent to the previous one, is the following: There are 2^m probabilities with sum *one*, $l(\epsilon_1, \epsilon_1, \cdots \epsilon_m)$ where ϵ_i equals either zero or one. $\epsilon_i = 1$ means that the transmitted gene corresponding to the ith factor is the maternal one while $\epsilon_i = 0$ means that the paternal value has been transmitted. Instead of (7) we have

$$l(\epsilon_1\epsilon_2 \cdots \epsilon_m) = l(1-\epsilon_1, 1-\epsilon_2, \cdots, 1-\epsilon_m)$$

These form a system of *quadratic difference equations with constant coefficients.* It has been shown that this solution can be found [7a]; it is however no simple matter so long as everything remains unspecified. In the case of independent assortment a very elegant solution has been given by H. Tietze [24] who established the recurrence relations for this case and investigated the limit behaviour in full generality. Jennings [13, b, c] and Robbins [20] have solved the general linkage problem for $m=2$, where one parameter enters, including an explicit expression for $p^{(n)}$. In the case $m=3$, where three parameters are involved, the explicit solution given by the author [7b] is still very simple.

Unless we have *complete linkage* ($l(1, \cdots, 1)=l(0, \cdots, 0)=\frac{1}{2}$, all other linkage probabilities zero) there is *no equilibrium for finite n.* The investigation of the *limit behaviour* of $p^{(n)}(z)$, as $n\to\infty$ becomes important. The author has established [7a] general results which may be formulated in the following statements:

First it is shown that, just as in the case $m=1$, *the gametic proportion of each single gene remains constant through the generations*

$$p^{(n)}(z_i) = p^{(0)}(z_i) \qquad (i = 1, 2, \cdots, m, n = 0, 1, \cdots). \tag{11}$$

This is the mathematical expression of the "immortality" of the genes. Next, consider the joint distribution $p^{(n)}(z)$. There are two important particular cases which have to be settled first. *Assume* $l(00 \cdots 0)=\frac{1}{2}$; *then all recombination values equal zero,* or, in other words, all "mixed" gametes, containing partly paternal and partly maternal genes, are a priori excluded. *We then have for all n:*

$$p^{(n)}(z_1 \cdots z_m) = p^{(0)}(z_1 \cdots z_m). \tag{12}$$

Completely linked genes act like one. The other extreme is *that no recombination value equals zero. Then it follows that*

$$\lim_{n\to\infty} p^{(n)}(z_1 \cdots z_m) = p^{(0)}(z_1) \cdots p^{(0)}(z_m). \tag{13}$$

In Mendel-Tietze's [24] case of independent assortment each $c_{ij}=\frac{1}{2}$, hence the above condition is satisfied and (13) holds.

Finally, if $c_{ij}>0$ does not hold for *all* pairs i, j, the linkage distribution degenerates in various ways into $t\leq m$ *groups of complete linkage.* By that we mean a set of genes $(i, j, \cdots, k)$ *within which no recombination takes place,* i.e. all corresponding c_{ij} are zero. For example if (1, 3, 6) is such a set, it can be shown that the *marginal distribution* $p_{136}^{(n)}$ *is preserved through the generations:* $p_{136}^{(n)}=p_{136}^{(0)}$ *holds for all n,* besides (11). To describe the limit behaviour completely we use the term:

maximal set of completely linked genes for a subset T of S containing *as many as possible of the m numbers with no recombination within the set.* Assume e.g. $m=8$, and that the group (1, 4) as well as the group (2, 3, 5, 6) are each completely linked, while the recombination values c_{12}, c_{17}, c_{18}, c_{27}, and c_{28} are known to be different from zero. Then (14), (2356), (7), (8) are the maximal sets of completely linked genes and, in an obvious notation:

$$\lim_{n\to\infty} p_{12\ldots8}^{(n)} = p_{14}^{(0)} p_{2356}^{(0)} p_7^{(0)} p_8^{(0)} \tag{14}$$

or, more generally [7, b]:

If $S_1, \cdots, S_t$ are the maximal sets of completely linked characters ($t \leqq m$), each containing at least one element, then

$$\lim_{n\to\infty} p_{12\ldots m}^{(n)} = p_{S_1}^{(0)} \cdots p_{S_t}^{(0)}. \tag{15}$$

The general theorem may be expressed in a more complete way:

Under the conditions of our investigation (m arbitrarily linked genes, random mating, etc.): *For the gene distribution in successive generations the original* (marginal) *distributions of each "maximal group" S_i are preserved: $p_{S_i}^{(n)} = p_{S_i}^{(0)}$, for all n; as $n\to\infty$, the joint distribution $p^{(n)}(z)$ of the m random variables approaches a limit where the different sets of genes are independently distributed.*

It is obvious that from these results corresponding results for the $w^{(n)}(x; y)$ follow easily, by means of the equation which is the analogue of (6).

It is biologically important and mathematically interesting to consider the extension where *different linkage distributions for males and females* are assumed. At the suggestion of S. Wright, who considered the problem for $m=2$ and $m=3$, the author has investigated this extension [7, d], assuming two linkage distributions l and l'. It turns out that the recurrence relations change, but not in an essential way: the arithmetic means of the values of the two linkage distributions play a decisive role. There appear now two different distributions of gametes corresponding to the two sexes, and simple recurrence relations hold for their arithmetical mean. As $n\to\infty$ the *two* gametic distributions approach *the same limit* which is independent of the linkage distributions and where the m genes are independently distributed.

6. THEORETICAL GENETICS OF AUTOPOLYPLOIDS

Thus far we have discussed the extension of Mendel's original ideas

to the case of linkage. A second remarkable extension has come with the discovery of *polysomic inheritance*, as opposed to normal or diploid inheritance. We shall discuss the important and interesting case of *autopolyploids*. This problem has been studied by numerous biologists. Among them we mention, because of the more theoretical aspect of their work, R. P. Gregory, H. J. Muller, J. B. S. Haldane [2], [9 e], S. Wright [26, c], K. Mather [16, b, c], [6], and R. A. Fisher [5, b, c, d].

In the preceding pages we assumed "chromosome segregation" as opposed to "chromatid segregation." It can be shown easily that for normal or diploid organisms, as considered so far, there is, mathematically, no difference between chromosome segregation and chromatid segregation. Thus the results of section 5 are general. This situation changes for polyploids, where there is a definite difference between the two modes of segregation. Chromosome segregation may be considered as an approximation to chromatid segregation; this latter one actually prevails according to modern studies. Nevertheless we shall discuss mainly chromosome segregation, since even under this simpler assumption the problem of polyploids appears rather complicated and unfamiliar to the statistician.

In the simplest case where only *one* locus is considered, an autopolyploid organism may be described as follows: In case of a $2s$ ploid each gamete consists not of *one* but of $s \geqq 1$ genes. A genotype possesses *two sets, each of s genes*, each gene being represented by one of the r numbers $1, 2, \cdots, r$, the r alleles, where r may be greater than, less than, or equal to s. Although a genotype is mathematically described by two sets, each of s numbers, there are less than $N = r^{2s}$ distinct types since many types have to be considered as equal. First, we have as before: $(x; y) = (y; x)$. Moreover, it is assumed with respect to the maternal (paternal) heritage that there is no difference between the various permutations of the s numbers which constitute this heritage. In accordance with this, a type is denoted, e.g. for $s = 5$, by $(a_1^2 a_2^3 : a_1 a_4 a_6 a_7^2)$. It is then easily seen that there are

$$(16) \qquad N_1 = \tfrac{1}{2}R(R+1) \text{ genotypes where } R = \binom{r+s-1}{s}.$$

With respect to segregation the results of observations suggest the following: In the formation of a new individual each parent transmits to the offspring a set of s genes out of the $2s$ genes the parent possesses. The selection of these s genes happens according to a probability dis-

tribution called the segregation distribution, which need not be the same for males and females. We assume here, however, that there is one and the same segregation distribution for both sexes. Moreover, the segregation distribution is independent of the genotype of the parent and of n. Let us consider its definition. Out of $2s$ numbers a set of s numbers can be selected in S ways, where

$$S = \binom{2s}{s} = \frac{(2s)!}{s!s!}.$$

These S cases (there are six in case of the most often considered tetraploid, $s=2$) have been more or less tacitly assumed equally probable, first by Muller, then, as far as I know, by all biologists concerned with the problem. This assumption is not logically necessary and does not seem inevitable in the light of reported observations. We therefore introduce a segregation distribution where the S probabilities are not a priori assumed equal. On the other hand we must avoid introducing parameters that are not meaningful biologically. After considering the analogy to other linkage phenomena and studying the numerical results of observations, the author [7c] was led to the hypothesis that within the s transmitted genes the *proportion* of maternal and paternal genes plays a certain role. We thus make the following definition: Call

$$l_\alpha = \binom{s}{\alpha}\lambda_\alpha \qquad (\alpha = 0, 1, \cdots, s) \tag{17}$$

the probability that a specified set of α maternal genes be transmitted. Assume symmetry of paternal and maternal heritage:

$$l_\alpha = l_{s-\alpha}. \tag{18}$$

Since a set of α specified maternal genes may be combined with $(s-\alpha)$ paternal genes in $\binom{s}{\alpha}$ ways, we have

$$\sum_{\alpha}^{0\cdots s}\binom{s}{\alpha}l_\alpha = \sum_{\alpha}^{0\cdots s}\binom{s}{\alpha}^2\lambda_\alpha = 1. \tag{19}$$

Thus if $s=2\mu$ or $2\mu+1$, just μ parameters are introduced. These μ parameters may be partly or all equal to each other. If $\lambda_\alpha = 1/S$ we have "random chromosome segregation" (Muller, Haldane). Other particular assumptions may be considered.

Although there is a certain apparent similarity between this mathematical formulation and that in the linkage problem, the situations are in fact quite different. In the polyploid problem, any α maternal

genes may be segregated together with any $(s-\alpha)$ paternal ones, while in the linkage problem the new set of m genes has to contain exactly one value, either the x_i or the y_i, with respect to each of the m subscripts, $1, 2, \cdots, m$.

Our three basic distributions, the distributions of genotypes and of gametes and the segregation distribution have thus been defined. The next step, and the most important one, consists in finding a general recurrence formula which permits us to derive step by step the $p^{(n)}$ for $n=1, 2, \cdots$, starting with $p^{(0)}$. It turns out that there exists for polyploids too a surprisingly simple recurrence relation. It is based on the above concepts and on the use of certain types of "marginal distributions" whose definition is much less obvious than in the linkage problem; this recurrence law holds for any "r" and any "s." For a homozygotic gamete of type (A^s) our recurrence formula is simply

$$p^{(n)}(A^s) = \sum_{\alpha}^{0 \cdots s} l_\alpha \cdot \binom{s}{\alpha} p^{(n)}(A^\alpha) p^{(n)}(A^{s-\alpha}). \tag{20}$$

The $p^{(n)}(A^\alpha)$ and $p^{(n)}(A^{s-\alpha})$ are marginal distributions defined in agreement with the general concept of a marginal distribution. For example $p^{(n)}(A^\alpha)$ is the probability of a gamete with α A-genes, no matter what its remaining $(s-\alpha)$ genes may be. If we consider heterozygotes, r different allelomorphs, where $r \lessgtr s$, so that a gamete is of type $(a_1^{x_1} \cdots a_r^{x_r})$ with $x_1 + \cdots + x_r = s$, we obtain a result which is not much more complicated than (20).

By means of these recurrence relations the author [7c] has derived a *limit theorem*, as $n \to \infty$. Haldane [9e] has indicated a distribution which reproduces itself under random segregation and thus represents a state of equilibrium. This result leaves open the biological and mathematical question of whether such an equilibrium is actually reached, and if so, under what conditions. Our limit theorem may be formulated as follows:

Denote the r alleles by $a_1, \cdots, a_r$, *and consider one locus, and chromosome segregation. Then*

1) *As in case of a diploid* (sec. 1) *the gametic proportions corresponding to each single allele remain unchanged through the generations,* $p^{(n)}(a_i) = p^{(0)}(a_i)$, $(i=1, \cdots, r, n=0, 1, \cdots,)$.

2) *If and only if* $l_0 < \frac{1}{2}$ *(i.e. if "mixed gametes" are not a priori excluded) the joint distribution of gametes converges towards a limit where the alleles are independently distributed.*

$$\lim_{n \to \infty} p^{(n)}(a_1^{x_1} \cdots a_r^{x_r}) = [p^{(0)}(a_1)]^{x_1} \cdots [p^{(0)}(a_r)]^{x_r}. \tag{21}$$

In the case of Haldane's "random segregation" (where all segregation probabilities are equal to each other), the condition of our limit theorem is satisfied and our result thus proves and completes his statement.

While chromatid segregation describes the biological situation more correctly than chromosome segregation it seems that this latter, with an appropriate general segregation distribution, presents a useful approximation. It would lead us too far to attempt an explanation of chromatid segregation. A few remarks must suffice. Chromosome segregation is a particular case of chromatid segregation, where certain probabilities, corresponding to "double reduction," are assumed to equal zero. A segregation distribution for chromatid segregation has been introduced by K. Mather and R. A. Fisher [6]. This basic approach has been generalized by the author [7, f] so as to contain as particular cases Fisher-Mather's segregation distribution [6], Haldane's "random chromatid segregation" [9, e], and my general chromosome segregation distribution [7, c]. These investigations are for $s=2$, 3, and 4. They include (according to the program outlined in section 3) *recurrence relations*; *solution* of those difference equations; and *limit-results*. It is worth mentioning that in this case the recurrence relations, as well as the limit theorems, are *essentially different*—and this in a very interesting way—from the results (20) and (21). While these last are of the same type as the corresponding results (9) and (13), the problems of chromatid segregation introduce an entirely new type of statistico-biological laws.

The next step leads to the study of *polyploids under consideration of linkage* (several loci). In a particular case (data from tetraploid primulas), DeWinton and Haldane [2] have proposed a linkage theory for tetraploids under chromosome segregation. A suggestively simple segregation distribution is introduced on the basis of evidence which indicated a certain "pairing" as prevailing in this plant. In a more recent paper Fisher [5, d] considers linkage under chromatid segregation. His paper is concerned with the enumeration of gametes and of genotypes and to a certain extent with the definition of a segregation distribution, problems which in this general form are far from easy.

The formulation of the problem of $2s$-ploids with m loci must include the problem of diploids for m loci (sections 4 and 5) as well as that of a $2s$-ploid for one locus (section 6). The general problem must reduce to these problems and respective results if either $s=1$, or $m=1$. Further, the case of a $2s_1$-ploid with m_1 loci ($s_1 \leqq s$, $m_1 \leqq m$) must appear as a "marginal" case of the more general one. Accordingly, I introduce [7, e]

a sufficiently general segregation distribution. Recurrence formulas are also derived and the limit behaviour of the characteristic distributions is completely investigated. The results are simple and complete. It seems that these structurally clear recurrence formulas describe relevant features of a rather general biological situation: $2s$-ploid, m loci, r alleles, random mating, chromosome segregation,—with Mendel's theory of heredity as the basis. A similar remark holds for the limit theorems [7f].

The study of *linkage of polyploids under chromatid segregation* seems to be very difficult. The aim is, of course, to study the problem by means of a segregation distribution which is well adapted to the available data and which satisfies the theoretical requirements. With such a distribution at the basis we want to obtain a fairly complete insight into the theoretical side of the problem, including recurrence relations, stability, limit behaviour, and rate of approach to equilibrium. The conclusions reached so far for $m=1$ leave no doubt that we have to expect an entirely new type of results.

7. SOME OPEN PROBLEMS

The mathematical description of polyploid linkage under chromatid segregation constitutes an example of an open problem. There are many open problems within the limits set by the title of this paper. Let us mention a few examples.

The genetics of autopolyploids should be investigated under the assumption of different segregation distributions for the two sexes. In the vast domain of selection problems the results are still incomplete in many instances, even from a purely mathematical point of view. As a very simple example consider the first of the selection problems described in section 2. Here, a general integration has not been given in case of arbitrary selection coefficients; the limit behaviour of the distributions, however, can be completely described (cf. in particular Haldane's [9d] basic work on selection). If we assume different selection coefficients for the two sexes, then, to the author's knowledge, not even the limit behaviour has been investigated for the general case. The generalization from two to r alleles, almost trivial in case of random mating, introduces considerable difficulties in selection problems. Linkage under differential-viability-selection as well as polysomic inheritance are not easy problems. There are similar open problems for the various "systems of mating" which involve some choice of the mate. These examples have been chosen such as not to require the explanation of new biological situations. Other mathematical problems arise in

connection with so-called "sex-linked" inheritance in contrast to "autosomal" inheritance which has been considered throughout this paper. If the assumption of "non-overlapping generations" (see section 1) is dropped we are faced with completely new and quite difficult problems.

The results reported here and the problems explained may appear, even to the mathematically-minded biologist, as rather formal and somehow remote from his field of interest. Modern genetics is an extremely complex science, with relations to various fields of knowledge, such as general biology, biophysics and biochemistry, general anthropology, physiology, and psychology. "Formal genetics" as considered in this paper represents only one side of a manysided problem. We need not however lose sight of the totality of a problem if we follow up thoroughly only one aspect. Modern physics began when Galileo performed in a scientific way some very simple experiments and described them in mathematical terms; modern genetics started with Mendel's exact observations of simple phenomena and their mathematical description. Today, in biology as well as in other branches of knowledge, the mathematical approach hardly needs justification. Nevertheless the biologist may sometimes wonder whether the generalizations which the mathematician uses are of any value to him. Each kind of approach has its inherent logic and must abide by its own inner laws. The mathematical approach to biological problems, by way of abstraction and generalization, will prove rewarding at present and in the future.

SELECTED REFERENCES

[1] F. Bernstein
(a) "Variations- und Erblichkeitsstatistik," *Handbuch der Vererbungswissenschaft*, Bd. 1, pp. 1–96.
(b) "Zusammenfassende Betrachtungen über die erblichen Blutstrukturen des Menschen," *Zs. f. Induktive Abstammungs u. vererbungslehre*, Vol. 37, (1925).

[2] D. De Winton and J. B. S. Haldane, "Linkage in the tetraploid Primula Sinensis," *Journal of Genetics*, Vol. 24 (1931), pp. 1–44.

[3] Th. Dobzhansky, *Genetics and the Origin of Species*, Columbia Univ. Press, (1938).

[4] Th. Dobzhansky and L. C. Dunn, *Heredity, Race, and Society*, Penguin Books, New York, 1946.

[5] R. A. Fisher
(a) *The Genetical Theory of Natural Selection*, Oxford Press, 1930.
(b) "The theoretical consequences of polyploid inheritance for the third style form of Lythrum salicaria," *Annals of Eugenics*, Vol. 11 (1941), pp. 31–39.
(c) "Allowance for double reduction in the calculation of genotype frequencies with polysomic inheritance," *Annals of Eugenics*, Vol. 12 (1944), pp. 169–171.

(d) "The theory of linkage in polysomic inheritance," *Phil. Trans. Roy. Soc. London*, B, 233 (1947), pp. 55–87.
(e) "A quantitative theory of genetic recombination and chiasma formation," *Biometrics*, Vol. 4 (1948), pp. 1–13.

[6] R. A. Fisher and K. Mather, "The inheritance of style length in Lythrum Salicaria," *Annals of Eugenics*, Vol. 12 (1943), pp. 1–23.

[7] H. Geiringer
(a) "On the probability theory of linkage in Mendelian heredity," *Annals of Mathematical Statistics*, Vol. 15 (1944), pp. 25–57.
(b) "Further remarks on linkage in Mendelian heredity," *Ibidem*, Vol. 16 (1945), pp. 390–398.
(c) "Contribution to the heredity theory of multivalents," *Journal of Mathematics and Physics*, Vol. 26 (1948), pp. 246–278.
(d) "On the mathematics of random mating in case of different recombination values for males and females," *Genetics*, 33 (1948), pp. 548–564.
(e) "Contribution to the linkage theory of autopolyploids," I and II, *Bulletin Mathem. Biophysics* Vol. 11, (1949), I, pp. 59–82, and II, pp. 197–219.
(f) "Chromatid segregation of tetraploids and hexaploids," forthcoming *Genetics*, 34 (1949).

[8] H. Geppert and S. Koller, *Erbmathematik*. Leipzig 1938.

[9] J. B. S. Haldane
(a) *The Causes of Evolution*, London, New York, 1932.
(b) *New Paths in Genetics*, London, 1941.
(c) "The combination of linkage values and the calculation of distances between the loci of linked factors," *Journal of Genetics*, Vol. 8 (1919), pp. 299–308.
(d) "A mathematical theory of natural and artificial selection, *Transactions of the Cambridge Philosophical Society*, Vol. 23 (1924), pp. 19–41. Successive parts in later volumes: *Proceedings, Cambridge Philosophical Society*, 23 (1927), pp. 363–372, 607–615, 838–848, *ibidem*, 26 (1930), pp. 220–230, 27 (1931), pp. 131–142, 28 (1932), pp. 244–248.
(e) "Theoretical genetics of autopolyploids," *Journal of Genetics*, Vol. 22 (1930), pp. 359–372.
(f) "The rate of spontaneous mutation of a human gene," *Journal of Genetics*, Vol. 31 (1935).
(g) "Some problems of mathematical biology," *Jour. of Mathematics and Physics*, Vol. 14 (1935), pp. 125–136.

[10] G. H. Hardy, "Mendelian proportions in a mixed population," *Science*, Vol. 28 (1908), pp. 49–50.

[11] A. Heilbronn and K. Kosswig, "Principia Genetica," *The Journal of Unified Science*, Vol. 8 (1939), pp. 229–255.

[12] L. Hogben, *An Introduction to Mathematical Genetics*, W. W. Norton, New York, 1946.

[13] H. S. Jennings
(a) "The numerical results of diverse systems of breeding," *Genetics*, Vol. 1 (1916), pp. 53–89.
(b) "The numerical results of diverse systems of breeding with respect to two pairs of characters etc. etc.," *Genetics*, Vol. 2 (1917), pp. 97–154.

(c) "The numerical relations in the crossing over of the genes etc., etc.," *Genetics*, Vol. 8 (1923), pp. 393–457.

[14] W. Johannsen, *Elemente der exacten Erblichkeitslehre*. 2. Auflage, Jena, 1913.

[15] V. A. Kostitzin, *Biologie Mathématique*. (Préface de Vito Volterra) Paris, 1937.

[16] K. Mather

(a) *The Measurement of Linkage in Heredity*. London, 1938.

(b) "Reductional and equational separation of the chromosomes in bivalents and multivalents," *Journal of Genetics*, Vol. 30 (1935), pp. 53–77.

(c) "Segregation and linkage in autotetraploids," *Ibidem*, Vol. 32 (1936), pp. 287–314.

[17] G. Mendel, "Versuche über Pflanzenhybriden," *Verhandlg. d. Naturforschenden Ver. in Brünn*. Abhandlungen, IV. Bd., Brünn, 1866, pp. 3–47.

[18] T. H. Morgan, *The Theory of the Gene*, New Haven, 1928.

[19] K. Pearson, "On a generalized theory of alternative inheritance with special reference to Mendel's law," *Phil. Trans. Roy. Soc.* (A), Vol. 203 (1904), pp. 53–86.

[20] R. B. Robbins, "Applications of mathematics to breeding problems II, and III," *Genetics*, Vol. 3 (1918), pp. 73–92 and pp. 375–389.

[21] F. W. Sansome, "Chromatid segregation in Solanum Lycopersicum, *Journal of Genetics*, Vol. 27 (1933), pp. 105–132.

[22] F. W. Sansome and J. Philp, *Recent Advances in Plant Genetics*, Sec. Ed., London, 1939.

[23] L. H. Snyder and C. W. Cotterman, "Studies on Human Inheritance XVII, Gene frequency analysis etc.," *Genetica*, Vol. 19 (1937), pp. 537–552

[24] H. Tietze, "Über das Schicksal gemischter Populationen nach den Mendelschen Vererbungsgesetzen," *Zs. Angew. Math. u. Mech.*, Bd. 3 (1923), pp. 362–393.

[25] W. Weinberg, "Über Vererbungsgesetze beim Menschen," *Zs. f. Indukt. Abstammungs- und Vererbungslehre*, Vol. 1 (1909), pp. 277–330.

[26] S. Wright

(a) "Systems of mating. I–V," *Genetics*, Vol. 6 (1921), pp. 111–178.

(b) "Evolution in Mendelian populations," *Genetics*, Vol. 16 (1931) pp. 97–159.

(c) "The distribution of gene frequencies in populations of autopolyploids," *Proc. Nat. Acad. Sciences*, Vol. 24 (1938), pp. 372–377.

(d) "Statistical genetics and evolution," *Bulletin of the American Math. Soc.*, Vol. 48 (1942), pp. 223–246.

(e) "Isolation by distance under diverse systems of mating," *Genetics*, Vol. 31 (1946), pp. 39–59.

(f) "On the roles of directed and random changes in gene frequency in the genetics of populations." *Evolution*, Vol. II, (1948), pp. 279–294.

THE FITTING OF LOGISTIC CURVES BY MEANS OF A NOMOGRAPH

EUGENE A. RASOR
Federal Security Agency, Social Security Administration

"Growth" curves, such as logistic curves, are widely used. This article discusses a previous paper on this subject and shows how simpler nomographs for fitting the logistic curve may be developed. First, a simpler nomograph of the same type as in the previous article is presented, and then a different type is developed which has the advantage of practically no computations, simplicity of appearance, and a wider range of values.

SPURR AND ARNOLD have shown how to fit a logistic curve quite simply by means of a nomograph.[1] This paper will first present a nomograph for finding the upper asymptote determined in the same fashion as did Spurr and Arnold, but somewhat simpler in having two of the scales parallel, and in addition taking in more values. Next, this paper will develop a nomograph of a different type, by means of which the logistic curve may be more readily and more accurately fitted to three equidistant values.

Chart 1 shows the nomograph for determining the upper limit of the logistic curve,

$$y = k/(1 + e^{a+bx}),$$

with the three scales being y_1/y_0, y_2/y_1, and k/y_0. It will be noted that the latter two scales are parallel and cover roughly the same values as did the Spurr-Arnold nomograph. However, the other scale shows a range of from 1.4 to 30.0, as contrasted with only 1.5 to 4.2 in the Spurr-Arnold nomograph. Thus, not only is the nomograph in Chart 1 more symmetrical, but also by utilizing the entire sheet it can cover more possible cases.

---

[1] William A. Spurr and David R. Arnold, "A Short-Cut Method of Fitting a Logistic Curve," *Journal of the American Statistical Association*, March 1948.

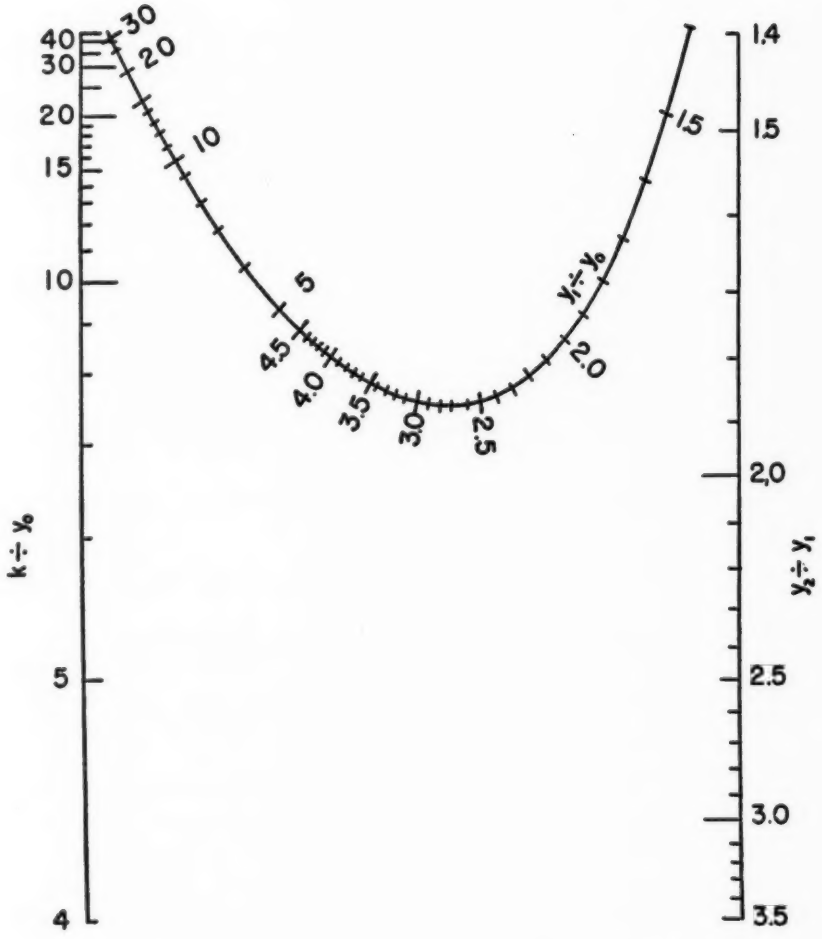


CHART 1. NOMOGRAPH FOR DETERMINING k IN $y = \frac{k}{1+e^{a+bx}}$

This nomograph has been determined from the following determinant:

$$\begin{vmatrix} 0 & \dfrac{44(\theta-4)}{41(\theta-1)}\mu & 1 \\ \delta & \dfrac{(7-2B)}{3B}\mu & 1 \\ \dfrac{396A\delta}{287A^2-178A+287} & \dfrac{44(7A^2-20A+28)}{287A^2-178A+287}\mu & 1 \end{vmatrix} = 0$$

where $A=y_1/y_0$, $B=y_2/y_1$ and $\theta=k/y_0$ as in the article by Spurr and Arnold; δ and μ are the width and height respectively of the chart.

Next, turning to the development of a simpler nomograph which will give accurate values of k, a, and b, directly from three equidistant values of y, without any preliminary calculations of the ratios of the y's and subsequent multiplication, along with intermediate (or extended) values of y_x directly from the graph, we have the following development:

For example, given the three equidistant values, y_0, y_1, and y_2, of

$$y_x = \frac{k}{1+e^{a+bx}},$$

let the problem be to find k, a, and b. First, plot the two points (y_0, y_1) and (y_1, y_2) on the nomograph, Chart 2. Then, draw a straight line connecting these two points and extend to intersect the 45° line, $y=x$. The x coordinate of this point of intersection is k. Connecting this point k with the point $(1+y_0, 1+y_0)$ and extending to the scale marked "$a+bx$" yields the value of a at the intersection. Similarly, connecting point k successively to points $(1+y_1, 1+y_1)$ and $(1+y_2, 1+y_2)$ yields $a+b$ and $a+2b$, respectively. The value of b is then determined by subtracting a from $a+b$; subtracting $a+b$ from $a+2b$ also gives a value of b and affords a check on the computations. Having determined a and b, we can now list the values of $a+bx$ for which y_x values are desired. The intersection with the 45° line of a line drawn between the point k on the y_x scale and $a+bx$ on the top scale yields $1+y_x$ on either the y_x scale or y_{x+1} scale from which y_x is readily determined by subtracting 1.

The mathematical solution of this problem in terms of y_0, y_1, and y_2 is

$$k = [2(1/y_1) - (1/y_0) - (1/y_2)]/[(1/y_1)^2 - (1/y_0)(1/y_2)]$$

$$e^{a+ib} = (k/y_i) - 1 \text{ where } i = 0, 1, 2, \text{ and}$$

where a and b are determinable from the last equation.

It can be shown[2] that if an equation $f(F, G, H)$ can be written as a determinant in the following form:

$$\begin{vmatrix} F_1 & F_2 & 1 \\ G_1 & G_2 & 1 \\ H_1 & H_2 & 1 \end{vmatrix} = 0$$

and three functions scales are drawn with coordinates $X_F = F_1$, $Y_F = F_2$, $X_G = G_1$, $Y_G = G_2$, $X_H = H_1$, $Y_H = H_2$ then the intersection of a straight line with these three scales yields a solution to the equation $f(F, G, H)$.

The above equation for k may be written in the following third order determinant:

$$\begin{vmatrix} 1/k & 1/k & 1 \\ 1/y_0 & 1/y_1 & 1 \\ 1/y_1 & 1/y_2 & 1 \end{vmatrix} = 0$$

which is the necessary and sufficient condition that the three points $(1/k, 1/k)$ $(1/y_0, 1/y_1)$ and $(1/y_1, 1/y_2)$ are collinear, when the function scales are determined as above. The above equation for e^{a+ib} may be written:

$$\begin{vmatrix} \delta/k & 0 & 1 \\ \delta/(1+y_x) & \mu/(1+y_x) & 1 \\ \delta e^{a+bx}/(1+e^{a+bx}) & \mu & 1 \end{vmatrix} = 0$$

which is the necessary and sufficient condition that the three points $(1/k, 0)$, $(1/1+y_x, 1/1+y_x)$ and $(e^{a+bx}/[1+e^{a+bx}], 1)$ are collinear, when the function scales are determined as before. Thus Chart 2 is basically double reciprocal paper and the scale at the top (from the above determinant) is $Y = \mu$ and $X = \delta e^{a+bx}/(1+e^{a+bx})$ where δ and μ are the total width and length of the chart grid, respectively.

The nomograph has five dotted lines showing the solution of the following specific problem: Given $y_0 = 2$, $y_1 = 4$, and $y_2 = 5$, to find k, a, b and some intermediate value of y_x, say, $y_{.5}$. First, the points (2, 4) and (4, 5) are plotted. The dotted line through these two points is extended to the line $y = x$. Then, reading the x coordinate of intersection on the y_x scale yields the value of k, namely, 5.3. The dotted line through the points (5.3, 0) and (3, 3) intersect the $a+bx$ scale at .51, which is the

[2] See Williamson, W. R. and Rasor, E. A., "Some Nomographic Theory and Applications to Benefits Under Retirement Plans," *Record of the American Institute of Actuaries*, XXX, 1941.

value of a. Likewise, the dotted line through the points (5.3, 0) and (5, 5) yields -1.10 for the value of $a+b$, and the dotted line through the points (5.3, 0) and (6, 6) yields -2.7 for the value of $a+2b$. By subtraction $(-1.10-.51)$, the value of b is -1.6, which is the same as

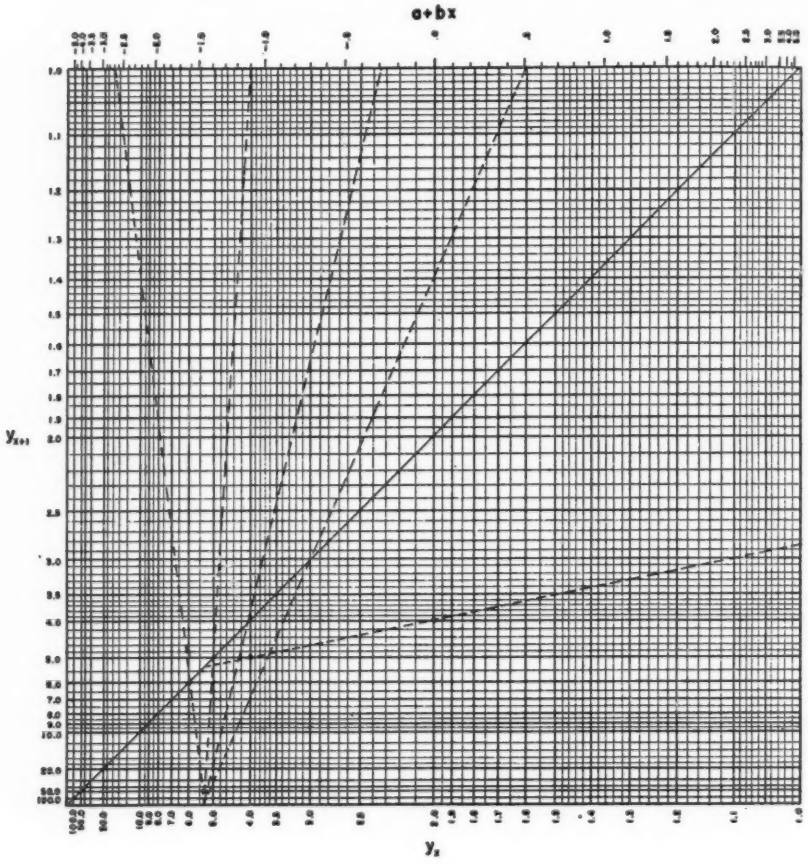


CHART 2. NOMOGRAPH FOR SOLUTION OF $y=\frac{k}{1+e^{a+bx}}$

obtained from the other subtraction, $-2.7-(-1.1)=-1.6$ so the computation is checked.

To find $y_{.5}$ we first determine $a+.5b$, which in this case is $.51+.5(-1.6)=-.29$ and the x or y coordinate of the intersection of the 45° line with the line drawn through 5.3 (the value of k) on the

y_x scale and $-.29$ on the $a+bx$ scale yields $1+y_{.5}=4.1$ or $y_{.5}=3.1$, the desired result.

In summary, it may be seen that the nomograph of Chart 2 has scales in simple linear form and may be entered directly without any preliminary calculations, while the result likewise is obtainable directly without any supplementary computations. A wide range is present for the various values of y, and if necessary this may be further broadened in three ways, by enlargement of the chart, a change of decimal place in all three given values of y, and by dividing the original data by y_0. These advantages indicate a much easier fitting of the logistic curve by a nomograph of the type of Chart 2, rather than one such as developed by Spurr and Arnold.

An even simpler nomograph for determining $1/k$ only would be to use a sheet of ordinary linear coordinate paper with a line drawn at 45° and the scales marked $1/y_x$ and $1/y_{x+1}$ in lieu of y_x and y_{x+1} shown in Chart 2. However, in using such a chart it would be necessary to predetermine the reciprocals of y and the result would yield the reciprocal of k. The scale for $a+bx$ could also be placed at the top as before, and resulting values for intermediate points would be $1/(y_x+1)$.

If more than three points are available, plotting all the points as above instead of a select three will show that the curve is of the logistic type if they lie on a straight line. On the other hand, if the points are not co-linear, an average value of k can be obtained from the chart.

ON THE BEST CHOICE OF SAMPLE SIZES FOR A t-TEST WHEN THE RATIO OF VARIANCES IS KNOWN

JOHN E. WALSH
The Rand Corporation

The situation considered is that of testing the difference of the means of two normal populations on the basis of a sample from each population, where the ratio of the population variances is known. The choice of sample sizes has been restricted to certain pairs which are equally preferable from the viewpoint of practical considerations (cost, difficulty of obtaining sample values, etc.). This note presents an easily applied method of determining which of these pairs of sample sizes yield the most powerful one-sided and symmetrical tests.

I. INTRODUCTION AND STATEMENT OF RESULTS

THE MOST powerful one-sided and symmetrical tests for comparing the difference of the means of two normal populations (variances unknown, ratio of variances known) with a given hypothetical value D_0 on the basis of sample values $x_1, \cdots, x_n$ from the first population and $y_1, \cdots, y_m$ from the second population are based on the t-statistic.

$$(1) \qquad \frac{(\bar{x} - \bar{y} - D_0)\sqrt{(n + m - 2)/(\theta/n + 1/m)}}{\sqrt{\frac{1}{\theta}\sum (x_i - \bar{x})^2 + \sum (y_i - \bar{y})^2}},$$

where θ equals the ratio of the variance of the first population to the variance of the second population (see [1]). This statistic has a t-distribution with $n+m-2$ degrees of freedom when the null hypothesis that the difference of means (mean of first population minus mean of second population) equals D_0 is true.

For the situation considered in this note, it is assumed that a number of pairs of sample sizes (n, m) have already been determined such that from the viewpoint of cost, inconvenience, etc., these pairs are equally preferable to the person applying the test; it is also assumed that the value of θ is known, either from past experience or by other means. The problem is to determine which of these given pairs of sample sizes yields the most powerful test (one-sided or symmetrical at the specified significance level) of the hypothesis considered. From above, the most powerful test will be based on the t-statistic (1) so

that this problem reduces to that of determining the pair of sample sizes which yield the most powerful test based on (1).

It is found (approximately) that the pair of sample sizes (n, m) which yields the most powerful one-sided test at significance level α and the most powerful symmetrical test at significance level 2α is that which furnishes the smallest value for the quantity

$$(\theta/n + 1/m)[1 - K_\alpha^2/2(n + m - 2)], \tag{2}$$

where K_α is the standardized normal deviate exceeded with probability α; i.e., K_α is defined by

$$\frac{1}{\sqrt{2\pi}} \int_{K_\alpha}^{\infty} e^{-x^2/2} dx = \alpha.$$

This criterion for choosing the pair of sample sizes which yields the most powerful test is reasonably accurate for $n+m \geqq 6$ if $\alpha=5\%$, $n+m \geqq 7$ if $\alpha=2.5\%$, $n+m \geqq 8$ if $\alpha=1\%$, $n+m \geqq 9$ if $\alpha=0.5\%$.

II. A FIELD OF APPLICATION

The results of this note are of most value when the ratio of the population variances can be considered known but the values of the variances are unknown. Some situations of this type are outlined below:

Let us consider a situation where the same treatment is applied to objects representative of populations of two different types. Then some common characteristic of the two types of objects is measured. The values thus obtained represent samples from each of the two types of populations, and these samples can be used to compare the relative effect of the treatment on the two types of objects (with respect to the specified characteristic). For example, the treatment might be a certain type of feed, the populations two different breeds of hogs, and the characteristic the weight at some specified future date.

Next let us consider the effect of different treatments on a fixed population. The mean and variance of the theoretical population of the values of the specified characteristic will vary with the treatment used. A not too uncommon situation is that where the variance of the theoretical population is a slowly increasing function of the mean of that population; i.e., if the magnitude of the observations increases, the variance of the observations also increases. Then a treatment which yields a theoretical population with a large mean value will yield a larger variance for this theoretical population than a treatment which results in a theoretical distribution with a smaller mean value. In the remainder of this section it will be assumed that the theoretical popu-

lations obtained for the specified characteristic have this property.

Finally let us consider applying a new treatment to two different populations of the type described above when data concerning the effect of some other treatment on these same populations are available. In many cases it can be assumed that the relative effect of the new treatment (with respect to the old one for which data are available) will be either to increase both of the theoretical population means or to decrease both of them. For example, an improved feed would be expected to increase the average weight of both breeds of hogs (though not necessarily the same amount) while a less satisfactory feed would be expected to decrease both average weights. If this is the case, the ratio of the variances of the theoretical distributions for the new treatment will tend to be about the same as the ratio of the corresponding variances for the old treatment; this ratio is therefore approximately known. Since the theoretical population means for the new treatment could be either larger or smaller than the corresponding means for the old treatment, however, estimation of the variances of the theoretical distribution for the new treatment from the corresponding values for the old treatment usually can not be done with any reasonable degree of accuracy. Thus the ratio of variances is available but the values of the variances are not.

III. EXAMPLE OF APPLICATION

Let us consider a case in which the only practical consideration is cost. Here the cost of an observation from the first population is always \$10, the cost of an observation from the second population is always \$100, while the total cost of the experiment is limited to \$400. Then the three pairs of sample sizes

$$(10, 3), \qquad (20, 2), \qquad (30, 1)$$

are equally preferable. Also for this particular situation it is known from past experience that $\theta=2$.

Let us determine which pair of sample sizes yields the most powerful symmetrical test at the 1% significance level. Then $\alpha=0.5\%$, $K_\alpha^2=6.62$ and the following table is obtained:

(n, m)	(2)
(10, 3)	0.76
(20, 2)	0.72
(30, 1)	1.20

Thus (20, 2) yields the most powerful symmetrical test at the 1% significance level.

A rough measure of the gain from using (20, 2) rather than (10, 3) is obtained by observing that (16, 2) furnishes approximately the same power as (10, 3). Thus about \$40 is lost by using (10, 3).

If the significance level of the symmetrical test had been 10% rather than 1%, the following table would have been obtained:

(n, m)	(2)
(10, 3)	0.61
(20, 2)	0.64
(30, 1)	1.12

Then (10, 3) would yield the most powerful test rather than (20, 2).

Since for fixed n and m the power of a one-sided or symmetrical t-test based on (1) increases monotonely as the value of

$$\sigma_1^2/n + \sigma_2^2/m = \sigma_2^2(\theta/n + 1/m)$$

decreases (σ_1^2 = variance of first population, σ_2^2 = variance of second population), it might be thought that the pair of sample sizes which furnishes the smallest value of $\theta/n+1/m$ would yield the most powerful test. That this is not necessarily so is seen by considering the above examples. The criterion based on $\theta/n+1/m$ is independent of the significance level while the two examples show that the choice of the pair of sample sizes which yields the most powerful test varies with the significance level.

IV. DERIVATIONS

This section presents proof of the statement that to a reasonable approximation the pair of sample sizes which yields the most powerful one-sided test at significance level α and the most powerful symmetrical test at significance level 2α is that which furnishes the smallest value for (2).

Let μ be the mean of the first population while ν is the mean of the second population. First consider the one-sided test of $\mu-\nu<D_0$ at significance level α and based on samples of sizes n and m respectively. Using a modification of the normal approximation given in [2], it is found that the power function of this type (1) t-test is approximately equal to

$$N\left\{-K_\alpha + \frac{(D_0 - \mu + \nu)}{\sqrt{\sigma_1^2/n + \sigma_2^2/m}}\left[1 - K_\alpha^2/2(n + m - 2)\right]^{1/2}\right\}$$

where by definition

$$N(z) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} e^{-x^2/2}dx.$$

This approximation is reasonably accurate for $n+m \geqq 6$ if $\alpha=5\%$, $n+m \geqq 7$ if $\alpha=2.5\%$, $n+m \geqq 8$ if $\alpha=1\%$, $n+m \geqq 9$ if $\alpha=0.5\%$.

Thus for $\mu-\nu<D_0$ and any permissible fixed values of α, σ_1^2, σ_2^2, D_0, μ, ν, the value of the power function is approximately largest when n and m are chosen so that (2) is as small as possible. This verifies the statement for the one-sided test of $\mu-\nu<D_0$. By symmetry the same result holds for the one-sided test of $\mu-\nu>D_0$.

Again using the modified normal approximation, the power function of the symmetrical type (1) t-test of $\mu-\nu \neq D_0$ is approximately equal to

$$2\alpha + \frac{1}{\sqrt{2\pi}}\int_{-K_\alpha}^{-K_\alpha+\delta} e^{-x^2/2}dx - \frac{1}{\sqrt{2\pi}}\int_{-K_\alpha-\delta}^{-K_\alpha} e^{-x^2/2}dx, \tag{3}$$

where

$$\delta = \frac{|D_0 - \mu + \nu|}{\sqrt{\sigma_1^2/n + \sigma_2^2/m}}\left[1 - K_\alpha^2/2(n+m-2)\right]^{1/2}.$$

Since $2\alpha<1$, $\alpha<0.5$ and (3) is a monotonely increasing function of δ. Thus for any fixed values of α, σ_1^2, σ_2^2, D_0, μ, ν the value of the power function is approximately greatest when n and m are chosen so that (2) is minimum. This verifies the statement for symmetrical tests.

REFERENCES

[1] Henry Scheffé, "On solutions of the Behrens-Fisher problem based on the t-distribution," *Annals of Math. Stat.*, Vol. 14 (1943), p. 43.

[2] N. L. Johnson and B. L. Welch, "Applications of the non-central t-distribution," *Biometrika*, Vol. 31 (1940), p. 376.

NOTE ON SOME ERRORS IN "THE EVIDENCE OF PERIODICITY IN SHORT TIME SERIES"

Several fundamental errors are in Truman Kelley's "The Evidence of Periodicity in Time Series." These involve a confusion of numbers of observations and degrees of freedom, the interpretation of probabilities obtained in tests of hypotheses and the appropriateness of periodigram analysis for detecting the existence of periodicity in time series.

I

THE ARTICLE by Truman L. Kelley, "The Evidence of Periodicity in Short Time Series"[1] contains several fundamental errors which should be noted. These particular errors are fairly widespread in popular literature and in practical applications. If these errors occur in this Journal and are not corrected their propagation will be insured since Kelley is deservedly a leading authority in Statistics. The younger generation, as well as the older, frequently rely on authority.

Originally, in 1943, the writer had considered it not worthwhile to write this corrective note. But recently on two occasions he has seen others using the technique presented by Kelley. On several other occasions during the war, interpretative errors to be mentioned below were made by persons who had taken formal university statistics courses and who were regular users of statistical methods. It is believed worthwhile, therefore, to point out these fundamental errors.

1. In testing goodness of fit of a trend line to a time series, Kelley treats the number of observations as independent observations in computing the number of degrees of freedom.[2] It is obvious (practically) that the observations are not independent.

2. In estimating and testing the period of a series of residuals, after removal of trend, the number of observations is again used in order to determine the number of degrees of freedom. In this particular analysis it might be contended that if periodicity does not exist in the residuals the residuals are independent; and hence, the assumption of independence is acceptable for purposes of testing the null hypothesis of non-periodicity. This contention would overlook the fact that the

[1] Truman L. Kelley, "The Evidence of Periodicity in Short Time Series," *Journal of the American Statistical Society* (1943), Vol. 43, pp. 319–326.

[2] *Ibid.*, pp. 320–325.

H. T. Davis, *The Analysis of Economic Time Series* (Bloomington, Indiana, 1941).

H. Wold, *A Study In the Analysis of Stationary Time Series* (Uppsala, Sweden, 1938).

Maurice G. Kendall, "On the Analysis of Oscillatory Time Series," *Journal of the Royal Statistical Society*, Vol. CVIII (1945), p. 93. This more recent article is cited because of its general excellence and comprehensiveness.

power of the test of non-periodicity as against various alternative lengths of periods is a function of the number of degrees of freedom (ignoring the extension into the sphere of varying amplitudes). As alternative hypotheses about the length of the period are admitted, the number of degrees of freedom applicable changes. In particular, the longer the period the fewer the degrees of freedom. If the degrees of freedom were constant, the test would have a greater probability of rejecting the non-periodicity hypothesis as the admissible alternative hypotheses of other period lengths are broadened to include longer periods. However, the degrees of freedom *do* decrease so that on balance if one assumed a fixed number of degrees of freedom he would not be controlling the probability of errors of the first type—namely rejecting the non-periodicity hypothesis when the series is in fact non-periodic.[3]

3. A common error of interpretation is present also. "Values of progressively less and less than .5 provide greater and greater evidence that b is not a chance deviation from zero."[4] By itself this statement is correct if it means that small values of P make it harder to accept the hypothesis that b is equal to zero (for a given power function). But it is apparent that other connotations are implied. "As P from this variance ratio is less and less than .5, there is greater and greater evidence that a period in the neighborhood of T exists."[5] If the test is biased there is no such evidence. The degree to which evidence is discriminatory is dependent upon the power function. In simple terms, evidence favors one alternative hypothesis only if it is relatively more probable under it and relatively less probable under other hypotheses. The following statement appears a little later: " . . . $P=.048$. Having odds of 952 to 48 that -1877.9 is not a chance deviation from zero. . . . "[6] The second sentence is a *non-sequitor*. If it were a valid extension, it would mean presumably that of all the times that one observes such deviations (those giving $P=.048$) .952 of such cases would in the long run have occurred from parameters not zero. Only by using Bayes' theorem can one obtain the type of statement employed by Kelley. The proportion of times in which such deviations are observed is a function of the experienced (true) parameters and not at all a result of the particular observed sample in any given case. What the probability does mean is that the probability of getting such a chance deviation or worse, *if* there is non-periodicity, is equal to .048. One *cannot* simply interchange

[3] J. Neyman and E. S. Pearson, "On the Problem of the Most Efficient Tests of Statistical Hypotheses," *Phil. Trans. Royal Society of London*, Series A, Vol. 231, p. 289.

[4] Kelley, *supra.*, p. 320.

[5] *Ibid.*, p. 323.

[6] *Ibid.*, p. 325.

the role of the hypothesis and the sample evidence and still have a correct statement.[7]

Later there appears "The smallest P is .152, found when $T=16$. Thus the odds are about 5 to 1 that a period in the neighborhood of 16 years exists."[8] In this statement and procedure of getting a P of .152 several errors are combined. a) The selection of the smallest probability from a set of independent probabilities will result in obtaining too many (more than would be obtained by a random selection) significant probabilities. This point has been covered by others.[9] It is especially applicable to periodigram analysis. b) The probabilities from the periodigram are not independent so that it is impossible to apply correct methods designed to test the significance of the largest (or smallest) of a set of independent probabilities. c) Finally the second sentence is again a *non-sequitor*. Suppose the probability had been .001. Would that then indicate that the odds are 999 to 1 that this is not a chance deviation and is instead a real difference? Obviously not. It merely says that *if* the truth is non-periodicity, then one would obtain by chance such a deviation only once in 1000 times (more accurately with probability .001). One cannot interchange the role of hypothesis and sample evidence! Consider the other extreme probability; suppose the probability in the periodigram for a particular period is 1.00. Clearly that is not *proof* that such a period is not in the series! Not only is the second sentence a *non-sequitor*, it is wrong. If it were correct it would be necessary that of all the cases in which deviations with probability .15 were observed, .85 of *those* cases would ('on the average') have come from situations in which there really was periodicity of length $T=16$! This requires Bayes' theorem. There is no indication that Kelley has anywhere in mind any *a priori* parameter distribution. And if he did, the simple complement of the probability is not the correct one if one wishes to apply Bayes' theorem.[10] The appropriate formula is somewhat more complex.

In summary of this particular error, the confusion appears to arise basically from a failure to understand adequately the fundamental procedure by which one increases his knowledge on the basis of evidence. Always one must ask how probable is the evidence that has been observed, *if* a certain hypothesis or cause is the true one. He *must also* ask

[7] Neyman and Pearson, *supra*.

[8] Kelley, *supra*, p. 326.

[9] W. G. Cochran, "The Distribution of the Largest of a Set of Variances as a Fraction of Their Total," *Annals of Eugenics* (London), Vol. 11, p. 47.

[10] J. V. Uspensky, *Introduction to Mathematical Probability* (New York: McGraw-Hill, 1937), pp. 60–73.

the same question about the probability of the observed evidence under the assumption that it came from some *other* hypotheses. After comparing these probabilities of obtaining the evidence (not of the hypotheses!) he can venture to draw conclusions. Never can he draw a conclusion until he has so obtained these different probabilities. At worst, he must make some intuitive guesses about the relative probabilities under various possible hypotheses. Lacking even an intuitive guess he can draw absolutely no conclusion whatsoever! And unless one uses Bayes' theorem, he simply cannot proceed to the next desirable (but nevertheless impossible step) of making a statement about the probabilities of the hypotheses![11]

4. Finally, the basic procedure used to detect periodicity—the use of a periodigram—is inappropriate. The periodigram technique used is appropriate to determination of the length of the period if the series is known to contain a period. Kelley's problem however is that of determining whether or not one exists.[12]

The writer regrets the occasion for this note, but these same errors are altogether too frequently committed elsewhere. It is another example of the adage that a statistician probably spends more of his time telling what not to do rather than doing.

ARMEN A. ALCHIAN

UNIVERSITY OF CALIFORNIA, LOS ANGELES

II

I do consider the number of observations as the number of independent items of information, or the number of degrees of freedom. No matter how brief the time interval between observations, if they are not consequent to the same instrumental errors—including judgment errors if estimates have been made—I would say that there is justification for calling them independent measures. If there is some other standard I fail to grasp it. However, I do not see that this point is crucial to Alchian's further criticisms.

Before discussing the main issue, which is that of inverse probability, I will comment upon the criticism of my statement that "The smallest P is .152. . . . Thus the odds are about 5 to 1. . . . " The value $P = .152$ was taken from a table giving P values for 17 different periods, 2-years, 3-years, . . . 18-years. As the value in that table there is no act of selection, but the moment it is taken out of that table because it is the

[11] Neyman and Pearson, *supra*.
[12] M. G. Kendall, *supra*.

smallest and put in the text there is an act of selection, so I admit that my statement, correct in the table, is a misstatement in the text and that Alchian's point (a) is sound. I am not sure about his point (b), but if in place of "impossible to apply" he had written "impossible for Kelley to apply" I would fully subscribe to it. His point (c) is in part covered in the first of this paragraph and is otherwise approached in the following discussion of inverse probability.

In the universe of situations in which designated null hypotheses do not hold other null hypotheses will hold. One very large class of these will be those in which the shift of a single parameter will lead to a null hypothesis which now holds. True, this is just a sub-class of the universe of situations, but a sufficiently large sub-class that a statement of probability covering it will be useful in determining conduct. In fact it seems to the writer that for most, perhaps all, of the problems of life a more valid determiner of conduct is unavailable. To illustrate: We have a null hypothesis that the true mean is M_1. The actual sample mean is M and we find that $P = .048$. That is, if the true mean is M_1 a deviation as great as that found will, in the long run, occur in .048 of the samplings. Now consider the divergence from the sample mean M of the true mean $\tilde{M}$. All parameters except the mean being constant in many samples of different sorts all of which have a sample mean M the true means will diverge from M by an amount greater than $(M_1 - M)$ in .048 of the samples. Thus in this restricted realm inverse probability holds. (We may note in passing that this restricted realm may be far less restricted by making linear transformations of the original variable.)

Let us say that this probability is small enough to lead to a certain course of action. Nothing but time and further experience will prove whether the course of action is sound or not. A double risk has been involved: first the $P = .048$ instead of some smaller amount and second is the materiality of other parameters than the one studied. As a matter of logic both of these risks must be present.

To set up alternative hypotheses and determine which of the two is more acceptable does not change the situation. It sharpens the issue so far as P is concerned, but it may augment the other risk for we must ask if the two hypotheses were the best two to use in determining future action.

Though I subscribe to the soundness of Alchian's reasoning I hold that the use of inverse probability has much practical warrant and usefulness. I suspect that in deciding upon a course of action from given

data Alchian and I would reach the same conclusion, though we employ somewhat different logical processes en route.

Alchian's final point—that the periodogram cannot be used to detect periodicity—is indeed sweeping. I know of no logical approach to the question of existence independent of that of amount to which the thing exists. Dr. E. L. Thorndike's oft quoted dictum "If a thing exists it exists in some amount" has been most serviceable in the field of psychology. It should equally serve other fields.

I appreciate the fine tone and precision of Alchian's article and have the hope that we are really not as far apart as his criticisms imply.

TRUMAN L. KELLEY

WILLIAM LANE AUSTIN (1871–1949)
JAMES CLYDE CAPT (1888–1949)

Two former Directors of the Bureau of the Census, both valued members of the American Statistical Association, have passed away in recent months. William Lane Austin, Director of the Census from 1933 to 1941 and a Senior member of the Association, died in Greenville, Mississippi on October 10. James Clyde Capt, who succeeded Austin as Census Director in 1941 and remained Director until his resignation because of illness on August 10, 1949, died in Washington on August 30.

A review of the careers of these two men—alike in some respects and differing widely in others—brings into sharp focus some of the changes during recent decades in our concepts of the role of statisticians in government service. Austin entered the temporary Census Office in 1900, two years before the Permanent Census Act created the present Bureau. He served successively as statistician in charge of the Census of Plantations; Chief Clerk; chief statistician in charge of the Censuses of Agriculture in 1920, 1925 and 1930; Assistant Director; and finally Director.

During this 40 year period he learned his statistics from experience with Census operations. And despite the brilliant analyses of census problems by General Francis A. Walker and some other former Directors, it was not generally recognized that Census operations required special skill or presented technical problems. Austin developed hard-headed shrewdness and skill as a negotiator and as the supervisor of personnel whose training in most cases, like his own, had been limited to experience in the ranks. It was when he became Director in 1933 that he first displayed abilities and attitudes that co-workers of a lifetime to quote the late Joseph H. Hill "had never suspected him to possess."

The immaturity that had previously characterized statistics as a profession was beginning to disappear and Austin felt the need to strengthen the personnel of the Bureau. With astonishing finesse in avoiding offense to life-long colleagues in the service—men who distrusted "impractical theorists"—he began building a staff having technical qualifications that he did not himself possess or even understand. Austin was the founder of the modern, efficient, and technically competent Bureau of the Census that serves the nation today.

Capt entered the Bureau of the Census in 1939, as Executive Assistant to Austin. Unlike the latter, he was without any previous experience in dealing with technical statistical problems, and was completely unknown to the statistical fraternity and to those users of data throughout the land who view the Census Bureau with somewhat the

same veneration that lawyers hold toward the Supreme Court. It is no exaggeration to say that many of these were profoundly apprehensive two years later when the still new Assistant was named Director, at Austin's retirement.

I well remember my own doubts en route to my first call upon the new Director, and the sense of reassurance with which I left his office. Nor was the spirit of trust and understanding engendered at that first contact ever dimmed in any of our later official or personal relations. Increasingly I came to regard "J.C.," as he was affectionately known, as a superb administrator. His direction of the Bureau's affairs impressed me as strict, unaffected by "old school ties," personal friendships or biases, and always fair and just. He was quick to admit mistakes and shortcomings, but coupled his admissions with determination that the same mistakes would not occur a second time. In his relations as Director of the Census with other agencies, he was always cooperative, unambiguous in his position, and decisive.

I believe, however, that the achievements of his administration sprang primarily from the quality that he shared so conspicuously with Austin, namely: respect for and deference to competence in others. This seems to me to have gone farther than the layman's frequent deferral to the specialist, and to have included a willingness to give his personnel free rein, even in areas of administration in which he could have claimed a superior right to judgment. They must prove their initiative by success, but the opportunity thus presented them reflected an inherent and commendable modesty in their chief. It made for staff loyalty.

The administrations of these two Directors represent, very fortunately, a consistent and continuous pattern of growth. The strength now exhibited by the Bureau of the Census, particularly in the high calibre of its technical and administrative staff, had its beginnings under Austin, and developed at an accclerated pace under Capt. The all-important need of the Bureau was the building up of the organization to a position of leadership in statistical techniques and operating methods. This seems to me the outstanding and culminating achievement of Capt's administration. Credit for his success must be shared with Philip M. Hauser, Ross Eckler, Howard Grieves and his other top assistants, upon whom the burdens of the Bureau administration during the crucial period of the Seventeenth Decennial Census now fall. Their tasks will be easier because of the solid process of Bureau-building that preceded.

STUART A. RICE
Assistant Director in
Charge of Statistical Standards
Bureau of the Budget

BOOK REVIEWS

(Dr. Oscar Buros' resignation as Review Editor was effective August 1, 1949. During the interim period until a new editorial committee is established, this section will be edited by Dr. Ernest Rubin for the Secretary's Office.)

Probability Theory for Statistical Methods. *F. N. David* (Lecturer in Statistics, University College, London, England). London and New York: Cambridge University Press, 1949. Pp. ix, 230. 15*s.*; $3.50.

REVIEW BY JOHN W. TUKEY

Associate Professor of Mathematics
Princeton University, Princeton, New Jersey

SIXTEEN chapters of from 11 to 18 pages each, and a short seventeenth, emphasize that this book is based on a carefully organized set of lectures. The imprints of Karl Pearson and Jerzy Neyman are plain to see. On the one hand, the introduction of the binomial distribution, in Chapter 3, is followed by derivations and applications of the evaluations of binomial probabilities with the Incomplete Beta Function Table, Uspensky's method using hypergeometric series, the normal approximation, and the Poisson distribution (which is regarded solely as a limit to the binomial). On the other hand, the Markoff theorem on least squares is carefully applied to stratified samples of minimum variance.

Noteworthy are (*i*) a 17-page chapter on simple genetical applications, (*ii*) the elementary derivation of sampling moments for sample moments from finite populations, (*iii*) an account of differences of zero and their simplest uses, (*iv*) an emphasis on mathematical precision and the use of elementary methods, and (*v*) a clear-headed middle-of-the-road attitude toward the relation of probability theory to the real world. While it is hard to name courses in the United States where such a book would be an appropriate text, it should be helpful supplementary reading for both students and teachers. (The reader is supposed to be familiar with the notation for third and fourth moments (p. 31).)

As is inevitable and appropriate, there are places where the author and the reviewer hold different views, but in only two places were definite errors detected. One is in the statement of the boundedness hypothesis for the Central Limit Theorem on page 217, where $0 \leqq m_2$ should read $0 < m_2$. The second is the derivation of Sheppard's corrections on page 214, where the cumulants of G are expressed in terms of those for E and x as if E and x were independent, when in fact E determines x completely.

The reviewer is unable to be sure of the meaning of "The fundamental probability set, written F.P.S. for short, will be just that set of individuals or units from which the probabilities are calculated" (p. 12) in view of the discussion on page 58. The only other place which seems likely to confuse

the student is in the discussion of the "elementary probability law" of a continuous random variable, where a careless reader could conclude that $p(x')$ was "the probability that x takes the value x'."

The author is definite on such points as replacing the binomial distribution by a continuous distribution with stepped densities, "Such an assumption is, of course, wholly fallacious, since the binomial probabilities are a discrete set of points, but it is a useful aid to memory if not pursued too tenaciously" (p. 53—more tenacity would have replaced npq by $npq+(1/12)$ with better approximation), the uselessness of the negative binomial, "It would appear wrong therefore to carry out calculations in which p and n are given negative values" (p. 65) and division by n or $n-1$, "If a measure of the scatter in the sample is required then—(dividing by n)—must be calculated. If it is desired to estimate the population standard deviation then the expression—(dividing by $n-1$)—may be calculated because in the long run it will be equal to σ" (p. 126). Two points need to be made in the last connection. First, if "in the long run" were well defined, the last statement would be wrong. Second, the reviewer, and, he believes, an increasing number of other statisticians, feel the time is coming soon to loosen the thrall of moments of inertia, and *always* divide by $n-1$.

References to further reading are given after each chapter. These could be slightly improved by giving dates (only 4 are given, including one on page 81) and by giving volume numbers in Arabic rather than Roman numerals (after all, statisticians no longer do arithmetic in Roman numerals).

After mentioning a few of the things he likes, and all of those he doesn't, it behooves the reviewer to point out what he failed to find. To fill the two prominent gaps in a reasonable manner, Miss David would have had to recast her order of presentation and bring in cumulants (semi-invariants) long before the next-to-last-chapter. If this had been done, and k-statistics had been added, then Irwin and Kendall's (*Ann. Eug.* 12: 138–142, 1944) powerful and simple methods of deriving sampling moments of sample moments from finite populations could have been used with a great gain in simplicity of presentation and in power of tools. (The experience of the early workers with heavy algebra and consequent errors points up the need for keen and powerful tools.) Second, the early introduction of cumulants would have made it possible to extend the present discussion of Lexis theory, which seems ineffective, into a presentation of the essentials of the analysis of variance *without normality assumptions* as developed by Pitman and Welch (cf. *Biometrika*, Vol. 29). This would have been a most important addition.

Quality Control by Statistical Methods. *G. Herdan* (Lecturer in Statistics, Department of Preventive Medicine, University of Bristol, Bristol, England). Edinburgh, Scotland: Thomas Nelson & Sons, Ltd., 1948. Pp. xi, 251. 21*s*.

REVIEW BY PAUL PEACH
Associate Professor, Institute of Statistics
University of North Carolina, Raleigh, N. C.

THE internal evidence of Dr. Herdan's book proves that he has at least some acquaintance with a wide variety of statistical methods. In his eight chapters he discusses not only the usual material of quality control literature (charts for fraction defective, charts for variables, and single and double sampling plans based on attributes) but such less customary subjects as correlation, analysis of variance, and inverse probability. He devotes a whole chapter to the use of probability graph paper, a tool that certainly deserves to be introduced to industrial workers. His is, I believe, the first text on statistical quality control to mention the important notion of components of variance. He includes notes on the t and chi-square tests, with a table of t (but not χ^2, "in order not to swell the number of appendices inordinately"). He has a brief note about generating functions. There is perhaps not another book in the statistical field that touches upon so many topics in so little space.

Elementary books are presumably written for the instruction of students, the general idea being that beginners are expected to read the book and learn from it certain lore that shall be true, or useful, or at any rate acceptable to authority. The student who picks up Dr. Herdan's book with any such expectation had better be prepared to take a substantial consumer's risk. Naturally, no ordinary book of 250 pages can include adequate expositions of all the topics Dr. Herdan has introduced. I don't believe anybody will learn to fit the least squares straight line by studying pages 115–118. The Latin square on page 102 is almost void of explanation or context. The discussion of acceptance sampling is based mostly on the work of Dodge and Romig; but there is only half a sentence about the AOQL concept. The professional statistician learns from these pages that Dr. Herdan knows about linear regression and Latin squares; the beginner can learn, I fear, nothing.

The book suffers further from a neglect of common standards of literary craftsmanship. Both the writing and the typography are frequently obscure and confusing, and sometimes we find downright misstatements that are obviously the result, not of ignorance, but of carelessness. On pages 3 and 4 we find what purports to be the equation of the normal curve, but so carelessly set up that no novice could be expected to read it rightly. In the discussion of the test for the significance of the difference of two means, some of the formulas are right, some wrong. In the chapter on probability graph paper the probability density function and the ogive are persistently confounded.

Even in his literary references Dr. Harden fumbles. He uses for one chap-

ter the whimsical subtitle "The Rod of Moses"; the rod in question was Aaron's (Exodus 7: 8–12) and the allusion seems hard straining for a defective analogy. Among his American authorities he mentions one, called Cowder in one place and Crowder in another; the reference is to my esteemed colleague, Prof. Dudley J. Cowden. This latter error is in some respects typical. It could have been avoided with a minimum of care; people who know Dr. Cowden will penetrate the veil of obscurity; others will derive no information from the reference.

One wonders how Thomas Nelson & Sons came to publish this book, what authorities served as their referees, whether the comments of these authorities were transmitted to Dr. Herdan, and what was done with the galley and page proofs. Something must have been neglected somewhere. At all events, the resulting book has in my opinion no possibilities for the instruction of students.

Statistical Methods in Research. *Palmer O. Johnson* (Professor of Education, University of Minnesota, Minneapolis, Minn.). New York: Prentice-Hall, Inc. (70 Fifth Ave.) 1949. Pp. xviii, 377. $7.65.

REVIEW BY FREDERICK MOSTELLER

Associate Professor of Mathematical Statistics, Department of Social Relations Harvard University, Cambridge 38, Massachusetts

A GREAT many people have probably been intending to write a modern statistical textbook similar to Johnson's. The notion is to write a rather advanced book suitable for students in education and psychology, suitable in the sense that the emphasis shall not be on the mathematics, but upon methods which are useful in research problems the students will face. People intending to write such a book now have two alternatives: they can either relax and forget about it, or they can raise their sights a good deal higher than they needed to before the publication of Johnson's book.

The arrangement of the material resembles that of a handbook, which is rather appropriate in view of the title. This arrangement will facilitate use of this book by active research people who will find the form very convenient. The table of contents is helpful in this connection. The index looks impressive, but the only thing I looked for I was unable to find (see below). The methods used are described carefully, but succinctly, and illustrated by worked-out problems on real data. Methods given are heavily referenced and every effort is made to send the student to original sources. I do not have much hope that students will go to the sources, but do believe that teachers will be grateful for the references because not all who will want to use this book will be prepared to teach from it. This book can be enhanced by instructors able to expand on the material presented. The first 200 pages could easily be inflated to 400 with no padding.

Tables of the normal (a poor one), *t*, chi-square, *F*, and one of Nayer's

for testing differences in variance among several samples of the same size are supplied. Over the years one gets sadder about tables of the normal. Why is it that authors only provide half of one table? It is perfectly true that by a little arithmetic one can get what one wants from such a table, but tables are for economy in time and errors. As a start I would like several normal tables: 1) cumulated from the left, 2) cumulated from the right (these should include negative arguments as well as positive), 3) cumulated symmetrically from the center, 4) cumulated symmetrically from the tails. I would also like companion tables using probability as the argument and the deviation as the entry. If it is objected that this is pure laziness, the energetic may still do their computations from the original function. I hope some textbook writers and publishers will cooperate in this matter soon.

The book opens with a much better than usual discussion of the realm of statistics. To the discussion of general uses of statistics in economic research I would add the possibility of statistical or probability models. In the chapter on probability and likelihood the author is not especially careful about distinguishing between the notions of a probability limit (p. 20) and the usual concept of limit. Bayes Theorem and Maximum Likelihood are too briefly treated, but the latter is discussed more fully later. The book begins to open up in Chapter III on Sampling Distributions, discussing means, differences between means, variance, t, r, the z transformation for correlations, the relation between z and F. Discussion of testing statistical hypotheses follows smoothly, comparing fiducial and confidence limits (making confidence limits sound a little more difficult than necessary), likelihood ratio and sequential tests.

Standard procedures are given for testing means against standards with standard deviation known and unknown, for finite as well as infinite populations, testing differences, the Behrens-Fisher problem, the sign test, testing differences in percentages using t. There is a footnote on page 81 which bothers me in this connection—"The χ^2 test is an exact test for this problem." If no correction is made for continuity the stated formula is equivalent to χ^2, but neither test can be regarded as exact as I see the matter. Continuing with this chapter, methods are given for testing equal variability for two or more groups, the significance of differences for correlations and regression coefficients, 2×2 and larger tables, goodness of fit. I will not continue this parade of techniques, but mention that later chapters cover estimation, interval and point, normalized distributions, special devices when the data are nonnormal, sampling theory and practice. All this is in 200 pages. The second half of the book handles analysis of variance and covariance, and multiple regression, including the discriminant function.

The analysis of variance examples are carried out in detail, one example running through 14 pages. The discussion is better than most such. Not enough emphasis is put on the meaning and interpretation of interactions, not enough is said about the limitations of the methods. I doubt that educators are always interested in such questions as: Are the mean achievements

of students in the three grades equal, or are the school means on the reading test equal? This is not so much a shortcoming of the book, but a shortcoming of the method as it is now available. This author does give with no heading, and no reference in the index that I can find, Fisher's method for comparing two particular means selected from a set (p. 234).

Perhaps the best thing to say is that Prentice-Hall's advertisement for this book is correct in all particulars—except one, I doubt if the author did research with R. A. Fisher (sic), although the influence of R. A. is clear.

Psychological Statistics. *Quinn McNemar* (Professor of Psychology, Statistics, and Education, Stanford University, Stanford, California). New York: John Wiley & Sons, Inc. (440 Fourth Ave.) 1949. Pp. 364. $4.50.

REVIEW BY EDMUND CHURCHILL

Assistant Professor of Mathematics, Department of Mathematics
Antioch College, Yellow Springs, Ohio

PROFESSOR McNemar's book is described as covering "all the statistical techniques, except factor analysis, that are frequently useful in psychological research." Altho "frequently used" might be more accurate than "frequently useful," and there are omissions in either case, this book does cover the standard topics from descriptive statistics through correlation, confidence limits, chi-square, sampling, and the analysis of variance, plus a chapter on the too often neglected topic of analysis of covariance.

The book, unfortunately, falls short of the author's aim to "provide a concisely, yet clearly written textbook which will lead to an appreciation of the place of statistics in psychological research." Conciseness is far from apparent at many places in the book, and neither the writing nor the logic behind it are always clear. This reviewer believes strongly that the student will best gain an appreciation of the role of statistics if clear, meaningful, and interesting illustrations are given as each technic is introduced. In this respect this book is decidedly weak. At least a dozen of the technics or procedures discussed in the book are not illustrated at all. Some are illustrated by ficticious data, others by data of the type: two groups classified according to 5 response categories (what groups? what categories?), 13 pairs of scores, IQ's of 161 five-year old boys, scores made by 50 college men on the Brown spool packer, etc. Here was an excellent opportunity both to enliven the text and to point the way to sound research practices by the frequent use of data from well designed (and well described) research.

There is a fuzziness in the discussion at a number of points. The discussion of the null hypothesis (p. 223) as it relates to the difference between means is not only muddled but at variance with an earlier discussion. The hypothesis of equal means does *not* imply equality of the population variances nor does the hypothesis of equa. means and variances lead to the t test. It is

interesting to note that, at different places in the book, the same hypothesis of equal means does not involve equality of variances (p. 65), is said to imply equality of variances (p. 223), and is described as involving such equality as an assumption (p. 225). The author would have stayed within his purposes if he had not only avoided this confusion of hypotheses and assumptions, but had also included tests of the hypothesis of equal means and variances and of the hypothesis of equal means without the assumption of equal variances.

The spirit of the null hypothesis is well violated in a procedure for testing the quantity $(p_1-q_1)-(p_2-q_2)$ where the p_i and q_i are non-exhaustive proportions in independent samples. Standard errors are obtained for (p_i-q_i) on the basis of the null hypothesis: $p_i-q_i=0$ $(i=1, 2)$. These standard errors form the basis for testing the difference mentioned above, altho the hypothesis that this difference is zero is a far cry from that used in getting the standard errors. The concept of the quality of a statistical test or confidence limit is missing. Obviously, no technical treatment of this concept is feasible but the awareness that we use one test or another in a given situation because we believe it is less likely to lead to error and that no test is always the best one is surely a basic part of the understanding of statistical inference. There are a number of minor points about the content of the book which may be worth noting. The square root of s^2 is, as usual, incorrectly described as unbiased. The sequence of histograms of $(\frac{1}{2}+\frac{1}{2})^n$ with fixed base length is asserted to converge to the normal curve; the sequence actually converges to a straight line. Sheppard's and Yate's corrections are given without hint that they can worsen instead of improve matters; this is especially true when Yate's correction is applied, as McNemar recommends, to 2×1 tables. The labelling of McNemar's graphs and the stubs of his tables, are, in several cases, atrocious and his treatment of approximate data is on the same level. There is no hint that in stratified sampling we might use samples in which the proportions are not in the same ratio as in the population, altho in general the best design calls for different ratios in sample and population. The equating of stratified sampling with the quota method may well confuse the student who has understood the quota method to be the one used by Gallup.

Frequent use of elementary algebra is made to derive formulas or to point the direction of their derivation. Altho there are those who will object to this practice, it is undoubtedly sound and a practice to be encouraged. On the other hand, it is difficult to understand why calculus is introduced to half-derive the linear regression equations when the elementary algebraic derivations of these equations are so simple and straightforward. Incidentally, a little of this algebra applied to the chi-square formula on page 207 would show that the maximum value of C for $2\times n$ tables is $\sqrt{1/2}$ and hardly rates classification as "unknown" (p. 182). One can also demonstrate that the maximum value of C for any $m\times n$ table is the same as for an $m\times m$ or an $n\times n$ table, whichever is the smaller.

The discussion of test reliability presents clearly some of the difficulties in estimating reliability, but its advice to base estimates on parallel forms is often completely unrealistic. The construction and administration of a second test form, if possible, would often represent an extremely inefficient use of research facilities. Split-half methods are mentioned but nothing is said about the effect of speed on these methods. The useful, reasonably sound, and rapidly computed Kuder-Richardson no. 20 is unmentioned. The simple derivation of this formula by Jackson and Ferguson might well have been included here.

The importance of homoscedasticity is stressed thruout the section on correlation (and might well have been stressed more often in the section on analysis of variance) but McNemar gives no hint as to how one may test a set of data for this mouth-filling property. Nor is the warning that many of the technics discussed are applicable only to normal data accompanied by advice or reference to sources of advice for the non-normal case. Much of this reviewer's pleasure at finding a chapter on covariance analysis is offset by the omission of such things as the discriminant function and the equivalent T^2 test, more useful, to be sure, than used, and the growing body of rapidly computed "inefficient" statistics which can be of considerable value in the early stages of a research program.

The above remarks are not intended to imply that there is not much which is clear and sound in this book. But statistics has reached a point in its development where we can rightfully expect that beginning texts have the clarity and thoro soundness that we expect and find in a college algebra text and do not find in this text.

Quality Control in Production: A Machine-Shop Manual on the Statistical Method of Controlling Product Quality During Manufacture. *H. Rissek.* Foreword by Frank Gill. London: Sir Isaac Pitman & Sons, Ltd., 1947. Pp. xvii, 181. 21*s*.

REVIEW BY H. A. FREEMAN

Associate Professor of Statistics, Massachusetts Institute of Technology
Cambridge, Massachusetts

THE title and subtitle describe the purpose of this book. Its chapter headings indicate its content. They are: 1, Statistical Method and the Quality Problem; 2, What is Quality Control? What are its Advantages? 3, The Basic Principles of Quality Control; 4, Control Charts based on Dimensional Measurement; 5, Control Charts based on Counting Defectives; 6, The Organization of a Quality Control System. There are also an adequate set of tables, a fair index, a good bibliography, and a foreword by Sir Frank Gill, Past President of the Institution of Electrical Engineers.

It is hard for me to say if this book is better or worse than its many competitors. Its statistical level corresponds to that now standard among qual-

ity control manuals; it must therefore be rated on its effectiveness in educating factory personnel. On this, I can only guess; my guess would be that this is a superior book, neither as facile or extended as some American manuals, but one which shows both that the author has digested his experience in industrial quality control and that he knows how to write about it.

One novelty: A discussion of Tippett's dual control chart is included. In this plan, the sampling results of go and of no-go gauges are used separately, the difference between the two numbers defective being a good estimate of the dimensional average, the sum, a good estimate of dimensional variability.

Sampling Methods in Forestry and Range Management, Second Edition *F. X. Schumacher* (Professor of Forestry, Duke University, Durham, N. C.) and *R. A. Chapman* (United States Forest Service, Washington, D. C.). Duke University School of Forestry, Bulletin 7, Revised. Durham, N. C.: the School, June 1948. Pp. 222. Paper, $2.00; cloth, $2.50.

REVIEW BY WALTER H. MEYER

Professor of Forest Management
Yale School of Forestry, New Haven, Conn.

THE first edition of this valuable treatise appeared in January 1942. The newly published revision attests well to the thoroughness and exactness of its predecessor, since scarcely a correction was found necessary in text, formulas, tables, or figures. One new chapter has been added, titled "Double sampling of individuals in representative sampling of groups; systematic computations." The authors must again be commended for assembling in one volume the variety of sampling techniques most appropriate to the complex fields of forestry and range management and for suggesting many methods of approach to sampling procedures in cases where earlier and current "attempts to extract sampling error are more akin to the art of the conjurer than to scientific assay." The authors state their purpose to be that of encouraging the practicing forester or range technician to "acquire the art of planning—and executing—suitable sample procedures, such that (1) the real error may be assessed unambiguously; and (2) the best estimate is obtainable . . . consistent with the time and funds available for the sampling work." Mathematical derivation and proof is held to a minimum and involved technical terminology is restricted with the result that the text can be readily understood by the forester or range manager with only a preliminary knowledge of the statistical method. If any criticism is to be made of this book, it must be toward the restriction indicated by the title, for surely the methods advocated will find a far broader application than in the fields of forestry and range management alone.

The first part, consisting of two chapters, deals with the statistical background of sampling in its simpler aspects. The second part (Chaps. 3–7) deals with direct estimates by sampling and the third part (Chaps. 8–13) with in-

direct estimates through regression. A short appendix supplies some of the essential mathematical background.

Part II on direct estimates by sampling takes up in order typical sampling schemes, starting from the unrestricted random sampling in finite and infinite populations, then proceeds to stratified sampling, the simultaneous sampling of two or more populations, subsampling and finally the representative sampling of irregular blocks of known or unknown area. Each of these has its appropriate place in the field of forestry. The forester may regret the scant attention that is given to systematic sampling, to which he has been addicted ever since the beginning of the use of sampling procedures in the estimation of wood volumes and which he probably will never discard completely. He knows that systematic sampling, especially with the double stratification that he uses (one for unit areas, the other for forest types) gives good results and he is supported in this view by the few known investigations involving the checking of 100 per cent enumerations by various types of sampling schemes. These have shown that systematic sampling gives results of high precision, superior to random or stratified random sampling under at least several typical conditions, but a precision which cannot be proven on the basis of the data and under current lack of a suitable mechanism of analyzing systematic samples. Since the publication of the first edition of this book, several authors have investigated the case of systematic sampling, including Osborne, Yates, Madow and Madow, Finney and others, and have demonstrated its suitability under certain conditions. The advantages of systematic sampling in forestry are so many from the administrative and financial point of view (and this is probably true also of many other fields) that statisticians could well devote more time and effort in developing a suitable theory.

Part III deals with the particularly valuable tool of regression as an aid in indirect estimation. Starting with simple linear regression for cases where the independent variable is free from sampling error and where it is not, it leads to the utility of purposive and mechanical selection of samples in obtaining an efficient regression equation. Conditioned regressions and the use of weights in such regressions is followed by a short treatment of non-linear regressions. Regression in representative sampling is shown to be an effective device of correlating ocular estimates with measured values taken on part of the general sample area. There follows the new chapter of double sampling in the representative sampling of groups, a topic which is only briefly treated, but appears to be the introduction of a rather detailed discussion of the systematic computation of normal equations. The latter appears to be the real purpose of this new chapter. The final chapter points out certain practical aspects of sampling, including the definition of objectives; bias; size, shape and structure of sampling units; the character of the sample itself; sampling intensity; and allocation of costs. A final new section, which does not appear in the earlier version, handles the allocation of optimum sample size to different strata. This section is called for in view of recent develop-

ments in aerial photogrammetry, whereby an advance stratification of a forested area can be made on a map showing groups of timber of varying average volumes and variances.

Cybernetics. *Norbert Wiener* (Professor of Mathematics, Massachusetts Institute of Technology, Cambridge, Mass.). The Technology Press. New York 16: John Wiley & Sons, Inc. (440 Fourth Ave.) and Paris: Hermann et Cie, 1948. Pp. 194. $3.50.

REVIEW BY SEBASTIAN B. LITTAUER
Associate Professor of Industrial Engineering
Columbia University, New York 27, N. Y.

CYBERNETICS is the science of communication and control in machines and men. The title was coined from the Greek $\chi\upsilon\beta\varepsilon\rho\nu\acute{\eta}\tau\eta$ in order to give identity to a new set of concepts developed by the author and his associates and further integrated by the author into a unified discipline as presented in this work. It encompasses a variety of branches of science in a manner which convincingly demonstrates the vitality and fruitfulness of the point of view of unification in the sciences. Many problems in neurology, psychopathology and communication engineering are shown to have a common core of meaning in which the methods of mathematics, logic and statistics are intrinsic and unifying factors.

The scope of the book is such as to appeal to a wide audience, including not only the specialists familiar with one or more of the particular fields dealt with, but also the non-specialist who is seriously concerned with the social implications of achievements in cybernetics. For, although the presentation is in considerable part mathematical, there is sufficient general discussion to acquaint and alarm any reader with the aims and potentialities of cybernetics. Of this the author takes cognizance, and in the closing pages of the introduction—an intensely interesting document on the circumstances which motivated and influenced these researches—he warns of the possible consequences to man that may follow from the creation of sensitive automata which can replace not only the human arm but also, in its simple functions, the human brain. In this, in spite of the fact that in the chapter on Newtonian and Bergsonian Time he argues the essential similarity between the functioning of the living organism and the cybernetic mechanism, he enjoins upon us the moral responsibility not to identify the human being as a commodity whose value is determined by the market place.

There are two fundamental aspects of the present work which largely contributed to its development as a new and unified science. One centers around the concept of the message and transmission of information and the other stems from the identification of the problems incident to the development of this concept in the nervous system with those encountered in some mechanisms. The common phenomenon of undesired hunting as a re-

sult of excessive feedback is quite analogous to a pathological condition known as purpose tremor. The recognition of this relation led to the study of certain aspects of neurophysiology as a feed back cycle of information circulating from the nervous system to the muscles and reentering the nervous system through the sense organs. The possibilities inherent in research from this point of view were presented by Wiener, Rosenbluth, and Bigelow in a paper entitled *Behaviour, Purpose and Teleology* which appeared in Philosophy of Science in 1943. The similarity of the information cycle in man and mechanism is suggested as a fruitful means of study of the former by experiment with the latter.The promise of this work presents a strong argument for combined effort by workers in each of these fields, fortified by some common ground of knowledge and vocabulary. It is part of the mission of cybernetics to encourage this closer tie.

The message—to quote—"is a discrete or continuous sequence of measurable events distributed in time—precisely what is called a time series by the statisticians." An operator or apparatus—predicitor—which follows a message finds conflict between the necessity for fidelity of response to smooth inputs (slow rates of change) and rapid response to sudden and large changes of input. Reconciling this conflict requires an appeal to the statistics of time series and the calculus of variations in order to find an operator deemed one of optimum prediction in that the mean square error of prediction is minimized.

The problem of disentangling a message from contaminating noise, or for that matter separating two messages, encountered in wave filter design presents a similar statistical picture. The transmission of information is statistical in that a prediction operator, in some specified sense optimum, must be based on the statistics of the time series of the message to be followed. The transmission of information is the transmission of alternatives, where the unit of information is that transmitted as a single decision between two equally probable alternatives. The amount of information is expressed as the negative logarithm to the base two of the number of such decisions that are to be made in attaining a particular observation. This it can be seen is the negative of the measure of entropy whence information or message is identified with negative entropy or the degree of organization in a system. The system of statistics developed for handling these problems becomes that of knowledge which can be expressed in a binary system. The results of these methods in determining prediction operators have been applied to computing machines, wave filters, simulated nervous system information cycles and the like, and found to be practically effective.

The mathematical developments are given in the chapters: Groups and Statistical Mechanics; Time Series, Information and Communication; Feed-Back and Oscillation. Familiarity with Lebesgue integration, probability theory and Fourier analysis, as well as with electrical circuit theory is helpful in reading these chapters. Nevertheless, their essence is remarkably well presented in verbal form. The statistics presented are based on knowledge of

the complete past of a time series; prediction based on sampling the history of a time series has but recently begun to be developed by a number of workers.

Implications and consequences of the theory developed are numerous. One possibility in quantum mechanics is quite promising; for, granting a hypothesis on the state of cosmic noise and degrees of freedom of the system the author's theory of information implementing the concept of negative entropy may be the proper means for deriving the Schroedinger equations from the Maxwell equations. It will be interesting to observe developments in this direction.

A striking feature of this book is the highly provocative nature of some of the author's speculations. For example, anyone familiar with problems in psychiatry and related attempts at practical therapy will be strongly persuaded by the cybernetic explanation offered for the nature of a class of disturbance and for the occasional effectiveness of its current treatment. Certainly within the field itself no explanation has met acceptance, and the stimulating suggestions offered by the author are to be welcomed for their promise of effective and practical progress. To the largest body of readers of cybernetics, and to the readers of this journal in particular, the chapter, Information, Language and Society, speaks most directly. The author is not optimistic about the immediate usefulness of the methods of cybernetics in the resolution of the problems of society, as are some of his colleagues in the fields of anthropology and sociology who have been urging him to take off in their directions. But he does show quite simply and directly in terms of the ideas developed in this book that most of the apologetics for the state of society in which our literature abounds, can in the very nature of things be but abortive and sterile efforts. In a framework within which there is no statistical control and in a time scale in which there can be only short runs, there can be no science of society, if science means prediction. The direction for fruitful work in the study of society is definitely pointed. If the influence of this book, and in particular of the last chapter, is such as to bring about a more realistic consideration of the problems of human survival, it will have done well.

Cybernetics had to be written—not only for the formal science which it presents—but also for the stimulation and enlightenment it offers to the non-scientist who in his turn does influence the direction of scientific inquiry. This is a unique book in its content—but it is also exceptional for its style, which manifests at times an expressive grace and incisiveness which cannot but compel the enthusiastic attention of the reader. Since it is reasonably certain that there will be further editions, a few improvements will be welcomed, such as correction of a number of misprints in the mathematical and verbal text, and inclusion of an index, as well as more references to the literature. In its depth, breadth and unity, Cybernetics is a powerful and important work; the author is to be congratulated for bringing his scientific knowledge and insight to bear upon problems of profound concern to a wide audience.

Sampling Methods for Censuses and Surveys. *Frank Yates* (Head of the Department of Statistics, Rothamsted Experimental Station, Harpenden, Herts, England). London: Charles Griffin & Co., Ltd., 1949. Pp. xiv, 318. 24*s*. (New York: Hafner Publishing Co., Ltd. $6.00.)

REVIEW BY W. EDWARDS DEMING
Adviser in Sampling, Bureau of Budget
Washington

HERE at last is a real book on sampling. It is a pleasure to review such an outstanding contribution. In the hands of an expert a probability sample can be designed so that it meets pretty closely some prescribed tolerance of sampling error (such as $\pm 2\%$, $\pm 15\%$), with a desired probability, and at the lowest possible cost per unit amount of information. Moreover, regardless of any assumptions that went into the planning, any estimate made from the sample is accompanied by an irrefutable index of precision. All this is made clear. In modern sampling practice it is also considered desirable to procure ancillary information concerning various alternative procedures, such as a different version of the questionnaire, a different plan of hiring, training, or interviewing, or supervising. This ancillary information illuminates the results obtained from the main survey, and is obtained by carrying out simultaneous or supplementary samples in addition to the main survey. Such ancillary information is particularly desirable for large surveys and complete censuses, as is coming to be the practice.

A sufficient background of mathematical theory and practical experience in dealing with human populations, farms, agricultural production, inventories, business activity, and numerous physical materials, have now accumulated by which probability samples may now be designed in a great variety of materials and problems.

The techniques of sampling have advanced rapidly during the past few years, but the requisite knowledge and experience have been largely confined to an inner circle of masters who have trained numerous apprentices by the spoken word. This is an unhealthy state of affairs, because the demand for competent statisticians has been running further and further ahead of the supply. No book can ever replace the privilege of working with a master, but Yates, indeed one of the masters, has put about as much inspiration and guidance into a book as is possible. Unfortunately, there is no royal and easy road to the top, but statistical teaching centres will now have an authoritative book to use for studies in sampling. The book will be useful for instruction in internships, and the private study of innumerable struggling, mathematically-inclined statisticians in government offices throughout the world will now be much more effective.

Other study and teaching aids in advanced theories of sampling do eixst; for example, Thionet's *Méthodes Statistiques Modernes des Administrations*

Fédérales aux États-Unis and *A Chapter in Population Sampling* produced by the sampling staff of the Bureau of the Census in Washington, 1947; also Mahalanobis's unsurpassed article entitled "On Large Scale Sample Surveys," in the *Phil. Trans. Royal Soc.*, Vol. 231B, 1944; and Walter Hendricks's *Mathematics of Sampling* (Virginia Agricultural Experiment Station, 1948), but these productions have not the scope of Yates's book.

The very title of the book is a clever statement of the attainment of modern methods of sampling. The title suggests, as is true, that sampling methods are now used not only for occasional or periodic special surveys of various kinds, but actually for replacing, broadening, and calibrating the data heretofore reserved for complete censuses of population, agriculture, commerce, etc.

The fact is that sampling is the modern approach toward all kinds of data. Sampling is the art and science of acquiring whatever information is desired, at the lowest possible cost, and with an objective index of precision. For detailed and precise tables by small areas, the most economical sample may of course turn out to be a complete census.

A glance at the table of contents is sufficient to show the extremely broad coverage that the author has included. Space will not permit a list of the topics treated, but they include a discussion of the ordinary biases that are encountered and the ways in which modern sampling design avoids and corrects these biases. The book begins at the beginning—the definition of the sample unit or units that are to be used, and the building or acquisition of the frame. In accordance with the recommendations of the U. N. Sub-Commission on Statistical Sampling, the term *frame* denotes a clear and unambiguous listing or mapping of the sampling units. In multi-stage sampling a frame will be required for every sampling unit that falls into the sample, in preparation for the next stage of sampling. The author discusses the efficiencies of various ways of drawing various kinds and sizes of sampling units, and of calculating the estimates.

In modern sampling design, the formula and procedure by which the estimates are to be prepared are as important as the procedures of selection; in fact, the two together constitute the sample design. Yates gives a splendid treatment of various methods of computation, along with excellent discussions of the amount of labor that is saved or entailed through the use of constant and variable weighting factors. He gives a well-balanced treatment of ratio-estimates, two-stage sampling, stratified sampling, variable sampling fractions, for some of which he was required to develop new theory to fill in the gaps. The estimate of the gains arising from multiple stratification (Sec. 8.4), and of using a plan of partial replacement of sampling units in successive samples (Sec. 88) will illustrate the breadth of techniques that are covered. A simple cost-function is treated on pages 283ff. In teaching the book, it might be well to make clear to the students that there is a cost-function associated with every sample design, and that sample-design in the

modern sense means using whatever plan will produce the most information per dollar.

The illustrations are mostly drawn from the author's wide experience in surveys of agricultural and wood lands, but this is excellent, because the fundamental principles are the same regardless of the material sampled.

The reviewer is in full agreement with the statement on page 59 that where a full census is compulsory, a sample should also be compulsory.

The well-chosen bibliography at the back, divided into convenient sections, will be very useful.

This review can only conclude with thanks and congratulations to the author: his book heralds a turning-point in statistical history.

INDEX TO VOLUME 44: 1949

ARTICLES

ALCHIAN, ARMEN A., *Note on Some Errors in "The Evidence of Periodicity in Short Time Series"* 559

BERKSON, JOSEPH, *Minimum X^2 and Maximum Likelihood Solution in Terms of a Linear Transform, with Particular Reference to Bio-Assay* . 273

BERRETTONI, JULIO N., and GREB, DONALD J., *AOQL Single Sampling Plans from a Single Chart and Table* 62

BRENNAN, J. F., *Evaluation of Parameters in the Gompertz and Makeham Equations* 116

CHANDRA SEKAR, C., and DEMING, W. EDWARDS, *On a Method of Estimating Birth and Death Rates and the Extent of Registration* 101

CHEVRY, GABRIEL, *Control of a General Census by Means of an Area Sampling Method* 373

COCHRANE, D., and ORCUTT, G. H., *Application of Least Squares Regression to Relationships Containing Auto-Correlated Error Terms* 32

COCHRANE, DONALD, and ORCUTT, GUY H., *A Sampling Study of the Merits of Autoregressive and Reduced Form Transformations in Regression Analysis* 356

COHEN, A. C., JR., *On Estimating the Mean and Standard Deviation of Truncated Normal Distributions* 518

DEMING, W. EDWARDS, and CHANDRA SEKAR, C., *On a Method of Estimating Birth and Death Rates and the Extent of Registration* . 101

DODD, STUART C., *On Measuring Languages* 77

FREEDMAN, RONALD, and HAWLEY, AMOS, H., *Unemployment and Migration in the Depression (1930–1935)* 260

GEIRINGER, HILDA, *On Some Mathematical Problems Arising in the Development of Mendelian Genetics* 526

GREB, DONALD J., and BERRETTONI, JULIO N., *AOQL Single Sampling Plans from a Single Chart and Table* 62

HAWLEY, AMOS H., and FREEDMAN, RONALD, *Unemployment and Migration in the Depression (1930–1935)* 260

JESSEN, RAYMOND J., *Some Inadequacies of the Federal Censuses of Agriculture* 279

KISH, LESLIE, *A Procedure for Objective Respondent Selection Within the Household* 380

KNOWLER, LLOYD A., and OLDS, EDWIN G., *Teaching Statistical Quality Control for Town and Gown* 213

KUZNETS, SIMON, *Wesley Clair Mitchell, 1874–1948. An Appreciation* . 126

LAWRENCE, NORMAN, and SHRYOCK, HENRY S., JR., *The Current Status of State and Local Population Estimates in the Census Bureau* 157

LESTER, A. M., *The Edge Marking of Statistical Cards* 293

METROPOLIS, NICHOLAS, and ULAM, S., *The Monte Carlo Method* . . . 335

MORRISON, NATHAN, *By-Product Data and Forecasting in Unemployment Insurance* 397

MOSER, C. A., *The Use of Sampling in Great Britain* 231

MOSTELLER, FREDERICK, and TUKEY, JOHN W., *The Uses and Usefulness of Binomial Probability Paper* 174
MYERS, ROBERT J., *Beneficiary Statistics Under the Old-Age and Survivors Insurance Program and Some Possible Demographic Studies Based on these Data* 388
NOETHER, GOTTFRIED E., *Confidence Limits in the Non-Parametric Case* . 89
NOVICK, DAVID, and STEINER, GEORGE A., *The War Production Board's Statistical Reporting Experience, V and VI* 413
OLDS, EDWARD B., *The City Block as a Unit for Recording and Analyzing Urban Data* 485
OLDS, EDWIN G., and KNOWLER, LLOYD A., *Teaching Statistical Quality Control for Town and Gown* 213
ORCUTT, G. H., and COCHRANE, D., *Application of Least Squares Regression to Relationships Containing Auto-Correlated Error Terms* 32
ORCUTT, GUY H., and COCHRANE, DONALD. *A Sampling Study of the Merits of Autoregressive and Reduced Form Transformations in Regression Analysis* 356
POLITZ, ALFRED, and SIMMONS, WILLARD, *An Attempt to Get the "Not at Homes" into the Sample without Callbacks.* 9
RASOR, EUGENE A., *The Fitting of Logistic Curves by Means of a Nomograph* 548
RICE, STUART A., *William Lane Austin (1871–1949); James Clyde Capt (1888–1949)* 565
SHRYOCK, HENRY S., JR., and LAWRENCE, NORMAN, *The Current Status of State and Local Population Estimates in the Census Bureau.* 157
SIMMONS, WILLARD, and POLITZ, ALFRED, *An Attempt to Get the "Not at Homes" into the Sample without Callbacks.* 9
SNEDECOR, GEORGE W., *On a Unique Feature of Statistics* 1
STEINER, GEORGE A., and NOVICK, DAVID, *The War Production Board's Statistical Reporting Experience, V and VI* 413
TUKEY, JOHN W., and MOSTELLER, FREDERICK, *The Uses and Usefulness of Binomial Probability Paper* 174
ULAM S., and METROPOLIS, NICHOLAS, *The Monte Carlo Method* . . . 335
WALLIS, W. ALLEN, *Statistics of the Kinsey Report* 463
WALSH, JOHN E., *On the "Information" Lost by Using a t-Test When the Population Variance is Known* 122
WALSH, JOHN E., *Applications of Some Significance Tests for the Median Which are Valid Under Very General Conditions* 342
WALSH, JOHN E., *On the Best Choice of Sample Sizes for a t-Test when the Ratio of Variances is Known* 554
WATKINS, RALPH J., *Statistical Requirements for Economic Mobilization* . 406
WILLCOX, WALTER F., *Conrad Alexander Verrijn Stuart (1856–1948)* . . 295
WOOFTER, T. J., *The Relation of the Net Reproduction Rate to Other Fertility Measures* 501

BOOK REVIEWS

ABBOTT, J. C., and BENAC, T. J., *Principles of Counting and Probability* . . . . . HERBERT SOLOMON 133
ANDERSON, J. L., and DOW, J. B., *Actuarial Statistics: Vol. II, Construction of Mortality and Other Tables* . . . . . T. N. E. GREVILLE 311

BRATT, ELMER CLARK, *Business Cycles and Forecasting, Third Edition* . . . . . WILLIAM A. SPURR 134
BRITISH STANDARDS INSTITUTION, *Fraction-Defective Charts* . . . . . ALBERT H. BOWKER 132
CHAPIN, F. STUART, *Experimental Designs in Sociological Research* . . . . . MARGARET JARMAN HAGOOD 312
CHARLIER, C. V. L., *Elements of Mathematical Statistics* Including *Table of Poisson's Function by L. v. Borkiewicz* . . . BURTON H. CAMP 313
CHURCHMAN, C. WEST, *Theory of Experimental Inference* JOHN W. TUKEY 136
CROXTON, FREDERICK E., and COWDEN, DUDLEY J., *Practical Business Statistics, Second Edition* . . . . . ALFRED CAHEN 444
DAHLBERG, GUNNAR, *Mathematical Methods for Population Genetics* . . . . . HOWARD LEVENE 447
DAVID, F. N., *Probability Theory for Statistical Methods* . JOHN W. TUKEY 567
DOUGLASS, RAYMOND D., and ADAMS, DOUGLAS P., *Elements of Nomography* . . . . . JOSEPH ZUBIN 315
EMMENS, C. W., *Principles of Biological Assay* . . LILA F. KNUDSEN 448
ENRICK, NORBERT L., *Quality Control: A Manual of Quality Control Procedure Based Upon Scientific Principles and Simplified for Practical Application in Various Types of Manufacturing Plants* J. H. CURTISS 139
GALLUP, GEORGE, *A Guide to Public Opinion Polls, Second Edition* . . . . . ROBERT COBB MYERS 315
GREENSHIELDS, BRUCE D., SCHAPIRO, DONALD, AND ERICKSEN, ELROY L., *Traffic Performance at Urban Street Intersections* HARRY G. ROMIG 142
. . . . . HENRY K. EVANS 451
HALD, A., *The Decomposition of a Series of Observations Composed of a Trend, a Periodic Movement, and a Stochastic Variable* . . . . . D. B. DELURY and BOYD HARSHBARGER 317
HENDRICKS, WALTER A., *Mathematics of Sampling* . . T. A. BANCROFT 144
HERDAN, G., *Quality Control by Statistical Methods* . . . PAUL PEACH 569
HILL, A. BRADFORD, *Principles of Medical Statistics, Fourth Edition* . . . . . MARGARET MARTIN 146
JEFFREYS, HAROLD, *Theory of Probability, Second Edition* . . . . . HERBERT ROBBINS 453
JOHNSON, PALMER O., *Statistical Methods in Research* . . . . . FREDERICK MOSTELLER 570
KENDALL, MAURICE G., *Rank Correlation* . . . . E. J. G. PITMAN 454
KENNEDY, CLIFFORD W., *Quality Control Methods* . SEBASTIAN B. LITTAUER 320
. . . . . CHARLES R. SCOTT, JR. 322
KERRICH, J. E., *An Experimental Introduction to the Theory of Probability* . . . . . J. F. KENNEY 147
MAINLAND, DONALD, *Statistical Methods in Medical Research: I, Qualitative Statistics (Enumeration Data)* . . . . . JOHN W. FERTIG 148
MATÉRN, BERTIL, *Metoder at Uppskatta Noggrannheten vid Linje- och Provytetaxering. [Methods of Estimating the Accuracy of Line and Sample Plot Surveys]* . . . . . T. W. ANDERSON 323
MCNEMAR, QUINN, *Psychological Statistics* . . . EDMUND CHURCHILL 572
VON MISES, RICHARD, *Lecture Notes on Mathematical Theory of Probability and Statistics* . . . . . BENJAMIN EPSTEIN 326

PANSE, V. G., *Report on the Scheme for the Improvement of Agricultural Statistics* . . . . . S. LEE CRUMP 455
RASHEVSKY, N., *Mathematical Theory of Human Relations: An Approach to a Mathematical Biology of Social Phenomenon* . FREDERICK MOSTELLER 150
RICKER, WILLIAM E., *Methods of Estimating Vital Statistics of Fish Populations* . . . . . CHARLES M. MOTTLEY 456
RISSEK, H., *Quality Control in Production: A Machine-Shop Manual on the Statistical Method of Controlling Product Quality During Manufacture* . . . . . H. A. FREEMAN 574
SCHUMACHER, F. X., and CHAPMAN, R. A., *Sampling Methods in Forestry and Range Management, Second Edition* . . . WALTER H. MEYER 575
SUMNER, W. L., *Statistics in School*. . . . . F. G. CORNELL 327
VAN UVEN M. J., *Mathematical Treatment of the Results of Agricultural and Other Experiments, Second Edition*. . . . . G. A. BAKER 329
WIENER, NORBERT, *Cybernetics* . . . . SEBASTIAN B. LITTAUER 577
WILKS, S. S., *Elementary Statistical Analysis* . . . . T. A. BANCROFT 458
WINTNER, AUREL, *The Fourier Transforms of Probability Distributions* . . . . . J. WOLFOWITZ 330
WOLD, HERMAN, *Random Normal Deviates: 25,000 Items Compiled from Tract No. XXIV (M. G. Kendall and B. Babington Smith's Tables of Random Sampling Numbers)* . . . . . H. BURKE HORTON 331
YATES, FRANK, *Sampling Methods for Censuses and Surveys* . . . . . W. EDWARDS DEMING 580
ZEISEL, HANS, *Say It With Figures*. . . . . GREGOR SEBBA 332

American Statistical Association

Organized November 27, 1839
Incorporated 1841

The American Statistical Association is a scientific and educational organization. Its membership is not confined to professional statisticians but includes economists, business executives, research directors, government officials, university professors, and other persons who are seriously interested in the application of statistical methods to practical problems, in the development of more useful methods, and in the improvement of basic statistical data. Engineers, mathematicians, biologists, actuaries, sociologists, psychologists, and representatives of many other professions are included in the membership of the Association.

Regular membership	$ 8.00*
Student membership	4.00
Introductory membership (for the first dues payment of applicants under 30 years of age)	4.00
Members subscription to Biometrics	3.00
Associate membership in the Biometrics Section	4.00
Contributing membership	25.00

Subscription rate, $8.00 per annum. Prices for back issues available on request.

Additional information about the Association and membership application forms may be secured from the Secretary, 1603 K Street, N. W., Washington 6, D.C.

The Editors welcome the submission of articles and notes for possible publication in the JOURNAL. Manuscripts should be sent to the Editor, Journal of the American Statistical Association, 1603 K Street, N. W., Washington 6, D.C. (1) Send two or more copies of your manuscript (preferably in separate envelopes, to avoid danger of loss). (2) Leave one half of the first page of your manuscript blank, to be used for instructions to the printer. (3) Attach to the front of your manuscript a fifty word outline or summary. Authors who wish suggestions about the preparation of manuscripts and charts or information about editorial policies should address their inquiries to the Editor. Book reviews should be sent to the American Statistical Association, 1603 K Street, N. W., Washington 6, D.C.

---

* 5.00 of which is for one year's subscription to the *Journal* of the American Statistical Association, and 75 cents for one year's subscription to *The American Statistician.*

GEORGE BANTA PUBLISHING COMPANY, MENASHA, WISCONSIN